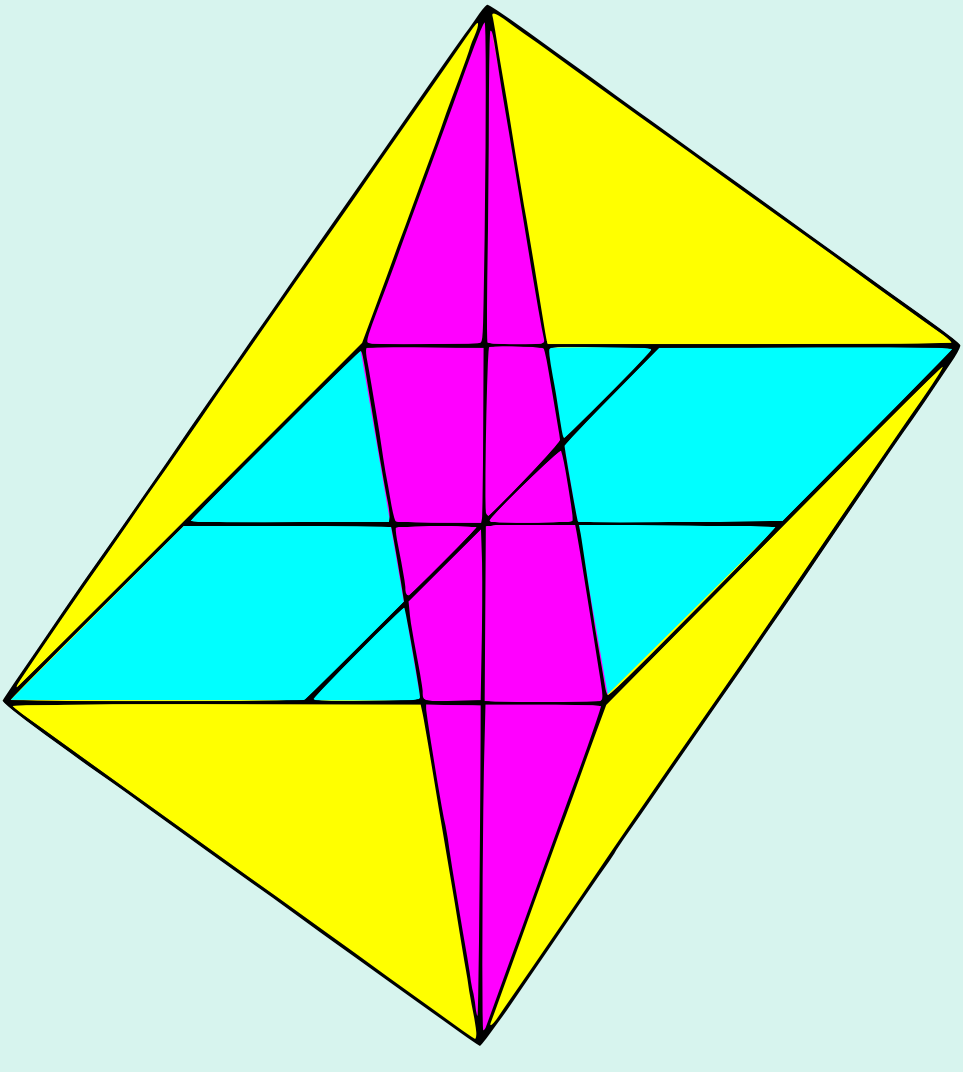
I.G. Petrovskii

PARTIAL DIFFERENTIAL EQUATIONS



PARTIAL DIFFERENTIAL EQUATIONS

Partial Differential Equations

I. G PETROVSKII

LONDON ILIFFE BOOKS LTD

Originally published in Moscow in 1961 by
Fizmatgiz under the title *Lektsii ob uravneniyakh s chastnymi proizvodnymi*

English edition first published in 1967 by
Iliffe Books Ltd., Dorset House, Stamford Street,
London S.E.1

Translated and prepared for press by
Scripta Technica Ltd

Printed in England by
Cox & Wyman Ltd.,
London, Fakenham and Reading

Contents

Preface

This standard work, which has already reached a third Russian edition, should be well-received in its translated form. The author is a mathematician of international reputation in the field of differential and integral equations. The present work deals mainly with the linear partial differential equations of mathematical physics, and should be a valuable text for mathematicians and mathematical physicists alike.

Chapter I

Introduction: Types of equations

1. DEFINITIONS. EXAMPLES

1. An equation containing partial derivatives of the unknown functions $u_1, u_2, \ldots, u_N$ is said to be an nth order equation if it contains at least one nth-order derivative but contains no derivative of order higher than the nth. The order of a system of partial differential equations is the highest order of any of the derivatives contained in any of the equations belonging to the system.

A partial differential equation is said to be linear if it is linear with respect to all the unknown functions and their derivatives that appear in it. A partial differential equation is quasilinear if it is linear with respect to all the highest-order derivatives of the unknown functions. Thus, the equation

$$\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x^2}+\frac{\partial u}{\partial y}\frac{\partial^2 u}{\partial y^2}+u^2=0$$

is a second-order quasilinear equation with respect to the unknown function u. The equation

$$\frac{\partial^2 u}{\partial x^2}+a(x, y)\frac{\partial^2 u}{\partial y^2}=2u$$

is a second-order linear equation with respect to the unknown function u. On the other hand, the equation

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2=u$$

is neither linear nor quasilinear with respect to that function.

A solution of a partial differential equation is any system of functions that, when substituted for the unknown functions in the equation, reduces the equation to an identity in the unknown variables. A solution of a system of equations is defined analogously.

In the present book, we shall deal primarily with linear second-order equations with one unknown function. The following are examples of such equations:

(1) $\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}$ (heat-flow equation);

(2) $\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}$ (the wave equation);

(3) $\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}=0$ (Laplace's equation).

Many physical problems lead to partial differential equations, including those just listed.

2. Example 1. The heat-flow equation Suppose that we have a body G whose temperature at the point (x_1, x_2, x_3) at the instant t is expressed by the function $u(t, x_1, x_2, x_3)$. We shall assume that the function $u(t, x_1, x_2, x_3)$ has continuous second-order derivatives with respect to the variables x_1, x_2, and x_3 and that it has a continuous derivative with respect to t.

The derivation of this equation, which describes the process of heat flow, is based on the following law. Suppose that a surface S is located within a body G and that a continuously varying vector n normal to this surface is defined at every point on it. The amount of heat q passing through the surface S to the side of the normal n in the instant of time from t_1 to t_2 is given by the formula

$$q=-\int_{t_1}^{t_2}\left\{\iint_S k(x_1, x_2, x_3)\frac{\partial u}{\partial n}\,dS\right\}dt. \tag{1.1}$$

Here, $\frac{\partial u}{\partial n}$ is the derivative at the point (x_1, x_2, x_3) on the surface S in the direction of the normal n. The inner integral is taken over the surface S.

The positive function $k(x_1, x_2, x_3)$ is called the coefficient of thermal conductivity of the body at the point (x_1, x_2, x_3).

Formula (1.1) is equivalent to the fact that an amount of heat equal to

$$dq = -k(x_1, x_2, x_3) \frac{\partial u}{\partial \cdot} dS\, dt.$$

passes through an infinitesimal area dS in an infinitesimal interval of time dt.

The physical law of thermal conductivity is usually expressed in this form.

If the area S lies on the boundary between the body and the surrounding medium, the following law holds. Suppose that $u(t, x_1, x_2, x_3)$ denotes the temperature of the body G at the point (x_1, x_2, x_3) just as before, and that $u_1(t, x_1, x_2, x_3)$ denotes the temperature at an arbitrary point (x_1, x_2, x_3) in the interior of the body. Then, the amount of heat entering the body through the area S on the boundary of the body in the time from t_1 to t_2 is given by the formula

$$q = \int_{t_1}^{t_2} \left\{ \iint_S k_1(x_1, x_2, x_3)(u_1 - u)\, dS \right\} dt, \tag{1.1'}$$

where the inner integral is taken over the surface. S. The functions u_1 and u are defined on S by passing to the limit from the outside and inside respectively. In this case $k_1(x_1, x_2, x_3)$ is the coefficient of thermal conductivity of the given medium (of which the body is made).

Let us consider a body that is isotropic with regard to thermal conductivity; that is, let us assume that the function $k(x_1, x_2, x_3)$ does not depend on the direction of the normal to the surface S at the point (x_1, x_2, x_3). In addition, let us assume that this function has continuous first derivatives with respect to all the coordinates.

To derive the heat-flow equation, let us pick out some subvolume D within the body G, that is bounded by a smooth surface S, and let us examine the change in the amount of heat within that volume during the interval from t_1 to t_2.

From formula (1.1), an amount of heat equal to

$$\int_{t_1}^{t_2} \left\{ \iint_S k(x_1, x_2, x_3) \frac{\partial u}{\partial n} dS \right\} dt, \tag{1.2}$$

where $\frac{\partial u}{\partial n}$ denotes the derivatives in the direction of the outer normal to the surface S, enters this subvolume.

On the other hand, this same amount of heat can be determined by the temperature change in the volume D during the interval from t_1 to t_2. The change in the amount of heat is equal to

$$\iiint_D c(x_1, x_2, x_3)\rho(x_1, x_2, x_3)[u(t_2, x_1, x_2, x_3) - u(t_1, x_1, x_2, x_3)]\, dx_1\, dx_2\, dx_3, \tag{1.3}$$

where $\rho(x_1, x_2, x_3)$ is the density and $c(x_1, x_2, x_3)$ is the specific heat of the body at the point (x_1, x_2, x_3), and the integral is taken over the region D*. When we equate (1.2) and (1.3) we obtain

$$\iiint_D c\rho\,[u(t_2, x_1, x_2, x_3) - u(t_1, x_1, x_2, x_3)]\, dx_1\, dx_2\, dx_3 = \int_{t_1}^{t_2} \left\{ \iint_S k(x_1, x_2, x_3) \frac{\partial u}{\partial n}\, dS \right\} dt. \tag{1.4}$$

By Ostrogradskii's formula,

$$\iint_S k \frac{\partial u}{\partial n}\, dS = \iiint_D \sum_{i=1}^{3} \frac{\partial}{\partial x_i}\left(k \frac{\partial u}{\partial x_i}\right) dx_1\, dx_2\, dx_3.$$

The integral on the left side of equation (1.4) can be written

* The values of the physical characteristics of a body at a particular point P (e.g. the density, specific heat, etc.) are always to be understood as a sort of limit. Specifically, take a sequence of cubes with centres at the point P such that the dimensions approach zero. For each cube, consider the ratio of the magnitude in question to the volume of that cube and take the limit of this ratio as the side of the cube tends to zero. Then, the density, for example, at the point means the limit of the ratio of the mass of the cube to its volume. Analogously, the surface density at a point on a plate means the limit of the ratio of the mass of a square with centre at that point to its area. The linear density at a point on a rod is the limit of the ratio of the mass of a segment with centre at that point to the length of the segment. Specific heat, thermal conductivity at a point, etc. are defined analogously.

in the form

$$\int_{t_1}^{t_2} \left\{ \iiint_D c\rho \frac{\partial u}{\partial t} \, dx_1 \, dx_2 \, dx_3 \right\} dt,$$

since

$$u(t_2, x_1, x_2, x_3) - u(t_1, x_1, x_2, x_3) = \int_{t_1}^{t_2} \frac{\partial u}{\partial t} \, dt.$$

Thus, for an arbitrary volume D within the body G, we have

$$\int_{t_1}^{t_2} \iiint_D c\rho \frac{\partial u}{\partial t} \, dx_1 \, dx_2 \, dx_3 \, dt - \int_{t_1}^{t_2} \iiint_D \sum_{i=1}^{3} \frac{\partial}{\partial x_i} \left(k \frac{\partial u}{\partial x_i} \right) dx_1 \, dx_2 \, dx_3 \, dt = 0$$

or

$$\int_{t_1}^{t_2} \iiint_D \left[c\rho \frac{\partial u}{\partial t} - \sum_{i=1}^{3} \frac{\partial}{\partial x_i} \left(k \frac{\partial u}{\partial x_i} \right) \right] dx_1 \, dx_2 \, dx_3 \, dt = 0.$$

Since the integrands are continuous and the volume D and the interval (t_1, t_2) are arbitrary, the following equation is valid for an arbitrary point (x_1, x_2, x_3) in the body G at an arbitrary instant of time t:

$$c\rho \frac{\partial u}{\partial t} = \sum_{i=1}^{3} \frac{\partial}{\partial x_i} \left(k \frac{\partial u}{\partial x_i} \right). \tag{1.5}$$

This equation is known as the heat-flow equation; it is valid in general for a nonhomogeneous but isotropic body. If the body is homogeneous, we have

$$k(x_1, x_2, x_3) = \text{const}, \quad c(x_1, x_2, x_3) = \text{const},$$
$$\rho(x_1, x_2, x_3) = \text{const}$$

and equation (1.5) is reduced to the equation

$$\frac{c\rho}{k} \frac{\partial u}{\partial t} = \sum_{i=1}^{3} \frac{\partial^2 u}{\partial x_i^2}. \tag{1.6}$$

If we set $\frac{k}{c\rho} t$ equal to t' and then denote this quantity by t

(instead of t'), we reduce the above equation to the form

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}. \tag{1.7}$$

Equations (1.5) and (1.7) have many solutions. To identify one particular solution from the entire set of solutions, we must have supplementary conditions, playing the same role as the initial conditions play in ordinary differential equations. Most often, supplementary conditions are what are known as boundary conditions, that is, conditions that are given on the boundary G of the space (x_1, x_2, x_3), for which the partial differential equation is applicable, and initial conditions applying to some particular instant of time.

It is physically clear, in the first place that knowing the temperature of the body at some instant of time and the temperature distribution on the boundary of the body must completely determine the temperature at any subsequent instant and, in the second place, that this temperature distribution can be given in a variety of ways. If the region G is all of space, we can show that a bounded solution of the heat-flow equation is uniquely determined for any $t > t_0$ by the initial conditions alone, that is, by the values of the function $u(t, x_1, x_2, x_3)$ at the instant $t=t_0$. For a bounded region G, it is possible, for example, to give the value of the temperature at every point of the body at some instant $t=t_0$ and, in addition, to give the value of the temperature at every point on the boundary of the body at all $t > t_0$. It turns out that these conditions are enough to determine uniquely a bounded solution for $t > t_0$ and $(x_1, x_2, x_3) \in G$.

Instead of giving $u(t, x_1, x_2, x_3)$ on the boundary of G for $t > t_0$, a unique solution to the heat-flow equation will be determined if we give the value of $\frac{\partial u}{\partial n}$ (i.e., the derivative of the unknown function u in the direction of the outer normal to the boundary of the region G) for points on the boundary of G. We arrive at such a mathematical problem if we study the temperature within the body G when we know the amount of heat passing from outer space to the surface of the body G through an arbitrary area S on the boundary of the body during an arbitrary instant (t_1, t_2) This amount must be equal to the amount of heat transmitted from the area S to the interior of the body. From formula (1.1), this quantity is equal to

$$\int_{t_1}^{t_2}\iint_S k\frac{\partial u}{\partial n}\,dS\,dt,$$

where $k>0$ is the thermal conductivity at the boundary point in question.

Thus, knowing the law for the transmission of heat for each area S on the boundary of the region G, we may find the value of $\frac{\partial u}{\partial n}$ on the boundary of G. In particular, if there is no heat exchange through the boundary, we have $\frac{\partial u}{\partial n}=0$ on the boundary.

Finally, as our boundary condition, the linear combination

$$k\frac{\partial u}{\partial n}+k_1 u,$$

where k_1 is the thermal conductivity upon transmission from the surrounding space to the body G and k is the thermal conductivity of the body, may be given on the boundary G for $t>t_0$. These coefficients are assumed known. We arrive at such a mathematical problem if we study the temperature within a body G when we know the temperature u_1 of the medium surrounding the body G. Then, in setting up a balance between the amount of heat passing through an arbitrary portion of the boundary G, we find from formulae (1.1) and (1.1') that (1) the amount of heat passing through the area S from the surrounding space to the surface of the body during an interval of time (t_1, t_2) is equal to

$$\int_{t_1}^{t_2}\iint_S k_1(u_1-u)\,dS\,dt.$$

(2) the amount of heat transmitted from an element of area S on the surface to the interior of the body during this same interval of time is equal to

$$\int_{t_1}^{t_2}\iint_S k\frac{\partial u}{\partial n}\,dS\,dt \qquad (k>0).$$

Since (t_1, t_2) and S are arbitrary, it follows that

$$k_1 u+k\frac{\partial u}{\partial n}=k_1 u_1.$$

In particular, if $u_1 \equiv 0$, this condition becomes

$$k \frac{\partial u}{\partial n} + k_1 u = 0.$$

Let us suppose that the temperature at each point (x_1, x_2, x_3) within the body G is constant with respect to time. Then $\frac{\partial u}{\partial t} = 0$ and equations (1.5) and (1.7) become respectively

$$\sum_{i=1}^{3} \frac{\partial}{\partial x_i}\left(k \frac{\partial u}{\partial x_i}\right) = 0, \quad \sum_{i=1}^{3} \frac{\partial^2 u}{\partial x_i^2} = 0. \tag{1.8}$$

This time, no initial conditions of any sort are necessary for determining $u(x_1, x_2, x_3)$. It will be sufficient to give just the boundary conditions, which must be independent of time. Physically, this can be easily interpreted as follows. If the boundary conditions do not depend on time, then, no matter what the initial temperature may be, the temperature $u(t, x_1, x_2, x_3)$ at every point (x_1, x_2, x_3) of the body approaches some limit $u(x_1, x_2, x_3)$ as $t \to \infty$. The limit function $u(x_1, x_2, x_3)$ satisfies the steady-state equations (1.8) and the original time-independent boundary conditions.

The problem of finding the solution of either of equations (1.8) from the values given on the boundary of the region in question is known as the Dirichlet problem or the first boundary-value problem.

In addition to the flow of heat in space, we often need to examine the change in temperature along a rod or in a plate. If the thickness of a homogeneous rod is such that the temperature can be assumed the same at all points of a single cross-section and if there is no heat exchange with the surrounding medium through the lateral surface of the rod, then the temperature u will depend only on the time t and one space coordinate x. In this case, the equation that the function $u(t, x)$ will satisfy will, with a suitable choice of units, be of the form

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}. \tag{1.9}$$

This equation would also be satisfied by the temperature $u(t, x_1, x_2, x_3)$ inside a three-dimensional body if this temperature depended on only one space coordinate, for example, on $x_1 = x$. This would be the case if the temperature of the body is the same at all points of each individual plane $x_1 = \text{const}$. Analogously, if we study the flow of heat in a

homogeneous thermally insulated flat plate, we arrive at the equation

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}. \tag{1.10}$$

3. Example 2. The equations of equilibrium and vibration of a membrane. A membrane is a stretched film that resists further stretching but does not resist any sort of distortion of its shape; that is, it does not resist any change in its shape that does not cause a change in the area of any portion of the membrane. The work of an external force that does cause a change in the area of some portion of it is proportional to that change. The positive proportionality constant T does not depend either on the shape of this portion whose area is changed or on its position. It is called the tension of the membrane.

We note for future reference that the work of the internal forces of elasticity is equal in absolute value to the work of the external forces that cause the change in area but has opposite sign.

Suppose that in a state of rest the membrane lies in the x_1, x_2 -plane and occupies some plane region G with boundary L.

Let us suppose that some force whose density at the point (x_1, x_2) is equal to $f(x_1, x_2)$ (see footnote following equation (1.3)) is acting on the membrane and that this force is directed perpendicularly to the x_1, x_2-plane. The membrane is bent under the action of this force and it takes the form of some surface whose equation we write in the form

$$u=u(x_1, x_2).$$

The u -axis is perpendicular to the x_1, x_2 -plane.

We shall devise an equation that the function $u(x_1, x_2)$ satisfies under the following restrictions. In the first place, we assume that the surface of the membrane is not bent excessively in the equilibrium position referred to above, that is, that it is almost a plane surface. In other words, we assume that the derivatives $\frac{\partial u}{\partial x_1}$ and $\frac{\partial u}{\partial x_2}$ are small, and in our calculations we shall neglect the higher powers of these derivatives. In the second place, we assume that, under the action of the force $f(x_1, x_2)$ the points of the membrane move only along lines perpendicular to the

x_1, x_2-plane, so that their x_1, x_2 -coordinates do not change.

Derivation of the equation will rest on one of the fundamental assumptions of mechanics, namely, the principle of possible displacements. According to this principle, at a state of equilibrium, the sum of the elements of work done by all forces acting on a system under any possible displacement (that is, allowed by the relationships that are imposed) must be equal to zero*.

To calculate the elements of work, we find the work done by forces acting on the membrane when it is displaced from its original plane position into the position given by the function $u(x_1, x_2)$. The work of a force whose density is $f(x_1, x_2)$, is given by the integral

$$\iint_G f(x_1, x_2)\, u(x_1, x_2)\, dx_1\, dx_2,$$

since a force of $f(x_1, x_2)\,dx_1\,dx_2$ acts on an element of the membrane $dx_1\,dx_2$. A change in the area of the membrane resulting from such a displacement is equal to

$$\iint_G \left(\sqrt{1 + u'^2_{x_1} + u'^2_{x_2}} - 1\right) dx_1\, dx_2,$$

and the work that the internal forces do when such a change in the area is made is equal to

$$-T\iint_G \left(\sqrt{1 + u'^2_{x_1} + u'^2_{x_2}} - 1\right) dx_1\, dx_2.$$

Let us expand this integrand in a series of powers of u'_{x_1} and u'_{x_2}. Remembering our assumption that these quantities are small, we discard the terms of higher powers in the expansion. Then, we get the following expression for the work done by the internal elastic forces

$$-\frac{T}{2}\iint_G \left[u'^2_{x_1} + u'^2_{x_2}\right] dx_1\, dx_2.$$

Therefore, the combined work**of all the forces acting on

* See SUSLOV, G.K., *Theoretical mechanics*, Gostekhizdat (1946).

**The integral (1.11) is equal in absolute value to the potential energy of the membrane in its equilibrium position. Thus, we may add that our derivation is based on the fact that the potential energy of any mechanical system has its minimum value in its equilibrium position.

the membrane and of the force f that results from displacement of the membrane from its position of rest into some position $u(x_1, x_2)$ is equal to

$$A(u)=\iint\limits_G\left[-\frac{T}{2}\left(u_{x_1}'^2+u_{x_2}'^2\right)+fu\right]dx_1\,dx_2. \qquad (1.11)$$

Let us now make some possible displacement of the membrane; that is, let us add to $u(x_1, x_2)$ some function $\delta u(x_1, x_2)$. The work that all the forces on the membrane perform when this displacement is made is equal to the variation of the integral (1.11), which is not difficult to compute. We have

$$\begin{aligned}\delta A&=A(u+\delta u)-A(u)\\&\approx\iint\limits_G\left[-T\left(u_{x_1}'\delta u_{x_1}'+u_{x_2}'\delta u_{x_2}'\right)+f\delta u\right]dx_1\,dx_2. \qquad (1.12)\end{aligned}$$

This variation must, by the principle of virtual work, be equal to zero.

Integrating the first two terms by parts, we obtain

$$\begin{aligned}&\iint\limits_G\left(u_{x_1}'\delta u_{x_1}'+u_{x_2}'\delta u_{x_2}'\right)dx_1\,dx_2\\&\qquad=\int\limits_L\frac{\partial u}{\partial n}\delta u\,ds-\iint\limits_G\left(\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}\right)\delta u\,dx_1\,dx_2,\end{aligned}$$

so that

$$\delta A=\int\limits_L -T\frac{\partial u}{\partial n}\delta u\,ds+\iint\limits_G(T\Delta u+f)\,\delta u\,dx_1\,dx_2, \qquad (1.13)$$

where Δu denotes the sum of the second derivatives $\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}$ and $\frac{\partial u}{\partial n}$ denotes the derivative in the direction of the outer normal to the boundary L. As was pointed out above, δu is a possible displacement, that is, a displacement allowed by the relationships imposed on the membrane. These relationships are usually stated for the edge of the membrane; therefore, the function $\delta u(x_1, x_2)$ for the interior points of the membrane is an arbitrary continuous function. Consequently, from the fact that δA is equal to zero, we may conclude that for the equilibrium position, the function $u(x_1, x_2)$ at any interior point satisfies the equation

$$T\Delta u+f=0. \qquad (1.14)$$

This equation is known as Poisson's equation.

The relationships referred to show up in the boundary conditions, and they may be of quite varied forms. We shall examine separately the cases that are most commonly encountered.

(a) ***The fastened membrane*** Suppose that the edge of the membrane is fastened along some space curve that is projected into L. Suppose that the parametric equations for L are $x_1 = x_1(s)$, $x_2 = x_2(s)$. Then, we require that the membrane pass through some curve $x_1 = x_1(s)$, $x_2 = x_2(s)$, $u = \varphi(s)$. In this case, the only restriction imposed on δu, is that $\delta u = 0$ on L. As a result of this restriction, the line integral in formula (1.13) disappears.

The problems that we have obtained, that of finding the solution to Poisson's equation with the boundary condition $u = \varphi(s)$ on L, is called the Dirichlet problem for that equation.

If $f = 0$ Poisson's equation reduces to Laplace's equation, which we have already encountered in the preceding example.

(b) ***A free membrane*** If we do not impose any restriction at all on the position of the membrane, its edge may move freely along the lateral surface of a cylinder whose base is L. In this case, δu is arbitrary both in the interior of the membrane and on the boundary G. We then obtain for equation (1.14) the condition on L

$$\frac{\partial u}{\partial n} = 0.$$

(c) The case is often encountered in which, in addition to the force f that acts on the interior points of the membrane, there is a vertical force of linear density f_1 acting on its edge. Therefore, a force $f_1 ds$ acts on an element ds of the boundary. If we seek the equilibrium position of the membrane under these conditions, we must add

$$\int_L \left(\int_0^u f_1 \, du \right) ds$$

and
$$\int_L f_1\, \delta u\, ds$$

to the integrals (1.11) and (1.12) respectively.

In this case, the line integral in formula (1.13) takes the form

$$\int_L \left(-T\frac{\partial u}{\partial n}+f_1\right)\delta u\, ds,$$

and we obtain for Poisson's equation a boundary condition of the form

$$T\frac{\partial u}{\partial n}-f_1=0$$

on the curve L. This problem is known as the second boundary-value problem (or Neumann's problem) if f_1 does not depend on u.

(d) Sometimes we need to consider the case of what is called 'elastic fastening of a membrane'. This is the case in which the force acting on the edge of the membrane is proportional to the displacement from the equilibrium position:

$$f_1(s)=ku(s).$$

Here, the boundary condition for Poisson's equation takes the form $T\frac{\partial u}{\partial n}-ku=0$.

Let us now turn to the derivation of the equation for the motion of a membrane. We shall consider only small transverse vibrations of the membrane. The smallness of the vibrations means that u, $\frac{\partial u}{\partial x_1}$ and $\frac{\partial u}{\partial x_2}$, are also small. We have already taken advantage of the smallness of the vibrations in a direction perpendicular to the x_1, x_2-plane. Thus, the coordinates (x_1, x_2) of a fixed point on the membrane do not change with change in t. The only change is in the value of the function

$$u=u(t, x_1, x_2).$$

The velocity of the point whose coordinates are (x_1, x_2) is

equal to $\frac{\partial u(t, x_1, x_2)}{\partial t}$. The acceleration is equal to $\frac{\partial^2 u(t, x_1, x_2)}{\partial t^2}$. To get the equation for the motion of the membrane, we need to calculate the initial force of the membrane according to d'Alembert's principle.

The density of this force is $-\rho(x_1, x_2)\frac{\partial^2 u}{\partial t^2}$, where $\rho(x_1, x_2)$ is the surface density of the membrane at the point (x_1, x_2). We shall obtain the equation for the transverse vibrations of the membrane if we replace the second term in equation (1.14) with $-\rho\frac{\partial^2 u}{\partial t^2}$:

$$T\left(\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}\right)-\rho\frac{\partial^2 u}{\partial t^2}=0. \tag{1.15}$$

The possible boundary conditions remain the same here as for equation (1.14) except that the given functions defined on the boundary may now depend on the time. Most frequently encountered are the problems of a membrane whose edge is fastened along a curve L (so that $u(t, x_1, x_2)=0$ on L, and the problem of the free membrane (so that) $\frac{\partial u(t, x_1, x_2)}{\partial n}=0$ on L.

As in the case of the heat-flow equation, it is physically obvious that the boundary conditions alone cannot determine uniquely the motion of the membrane since it depends largely on the initial position and velocity. In fact, it will be shown later that the solution of equation (1.15) is uniquely determined if we give the initial conditions

$$\left.\begin{aligned} u(t_0, x_1, x_2)&=\varphi_0(x_1, x_2),\\ u_t'(t_0, x_1, x_2)&=\varphi_1(x_1, x_2)\end{aligned}\right\}\quad (x_1, x_2)\in G \tag{1.16}$$

and boundary conditions of any of the types that we examined above.

Theoretically, we might consider the so-called unbounded membrane, that is, vibrations of the entire x_1, x_2-plane that obey equation (1.15). We shall encounter such a problem if the membrane in question is so large that we may neglect the effect of its boundary.

In such a case, as will be shown later, initial conditions are sufficient by themselves to determine the solution to equation (1.15) uniquely. On the other hand, if only conditions (1.16) are given for a finite membrane, this will de-

termine the solution uniquely not for all values of t but only for each point (x_1, x_2) in some interval $(-t_1, t_1)$, that depends on the point. Here, the closer the point (x_1, x_2) is to the boundary of the region G, the shorter will this interval be.

If ρ is constant, we can reduce equation (1.15) to the equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}. \tag{1.17}$$

by a change of the independent variables.

If we consider small vibrations of a gas (sounds waves) we may show that, under certain physical hypotheses, the function $u(t, x_1, x_2, x_3)$, which characterises the deviation from the normal pressure at the point (x_1, x_2, x_3) at the instant t, satisfies the equation

$$\frac{1}{a^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2}, \tag{1.18}$$

where $a > 0$ is a constant (the velocity of sound).

An equation of the form (1.18) is known as the wave equation in space. Many other vibrational processes (for example, electromagnetic processes) are also described by equation (1.18). Equation (1.17) is called the wave equation for a plane.

In the one-dimensional case (the vibration of a string, the vibration of a gas in a tube), the corresponding function u satisfies the equation

$$\rho \frac{\partial^2 u}{\partial t^2} = T \frac{\partial^2 u}{\partial x^2}. \tag{1.19}$$

This equation is known as the equation for a vibrating string. Here, $\rho(x)$ is the linear density at the point x and T is the tension in the string. The initial and boundary conditions for equations (1.18) and (1.19) are completely analogous to the corresponding conditions for equation (1.15).

We note again that equations (1.15), (1.18), and (1.19) can be obtained only if we neglect the quantities $\left(\frac{\partial u}{\partial x_i}\right)^4$ in comparison with $\left(\frac{\partial u}{\partial x_i}\right)^2$. If we do not do this (that is, if we do not assume that the vibrations are small in amplitude), the equations for the motion of the corresponding elastic bodies will be much more complicated and

they will not be linear equations.

Remarks 1: If we treat t as another space coordinate, the function $u(t, x_1, x_2)$ describing the vibrations of the membrane will be defined in a cylinder C whose generators are parallel to the t-axis and pass through the boundary of the region G over which the membrane is stretched. The problem that we examined above consisted in determining the values of this function within the cylinder from certain conditions given on its lateral surface and from the values of $u(t_0, x_1, x_2)$ and $u_t'(t_0, x_1, x_2)$ when the point $(x_1, x_2) \in G$ lies in the base of the cylinder. If we treat the problem this way, we no longer consider the initial conditions at $t = t_0$ as something significantly different from the boundary conditions. Both now become boundary conditions given on the boundary of the cylinder C.

2: When we considered the heat-flow equation and the equation for vibrations in an isotropic medium, these equations contained the expressions

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \quad \text{or} \quad \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2}. \tag{1.20}$$

These expressions are known as Laplacian operators (operating in this case on the function u) or simply Laplacians. They always appear in linear second-order two- or three-dimensional equations for a homogeneous isotropic medium because, up to a constant factor, they are the only linear combinations of the second partial derivatives of u that remain invariant under an arbitrary orthogonal transformation, that is, under a rotation of the orthogonal coordinate axes in two- or three-dimensional space.

2. THE CAUCHY PROBLEM. THE KOVALEVSKY THEOREM

1. Statement of the Cauchy problem. Consider the following system of partial differential equations in the unknown functions $u_1, u_2, \ldots, u_N$ of the independent variables $t, x_1, x_2, \ldots, x_n$

$$\frac{\partial^{n_i} u_i}{\partial t^{n_i}} = F_i\left(t, x_1, \ldots, x_n, u_1, \ldots, u_N, \ldots, \frac{\partial^k u_j}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}}, \ldots\right) \tag{2.1}$$

$$k_0 + k_1 + \ldots + k_n = k \leqslant n_j; \quad k_0 < n_j; \quad i, j = 1, 2, \ldots, N).$$

As one can see, for each of the unknown functions u_i, there is an integer n_i such that the highest-order derivative of u_i appearing in this system of equations is n_i. The independent variable t plays an especial role among the other independent variables because (1) among the derivatives of highest order (i.e., of order n_i) of each function u_i appearing in the given system is the derivative $\frac{\partial^{n_i} u_i}{\partial t^{n_i}}$ and (2) the system is solved explicitly for derivatives of this type. As a rule, t plays the role of time in physical problems and $x_1, x_2, \ldots, x_n$ are the space coordinates. The number of equations is equal to the number of unknown functions.

The values of the unknown functions u_i and their derivatives with respect to t of order $n_i - 1$ are given for some particular value $t = t_0$ (these are the 'initial values'). Suppose that, for $t = t_0$

$$\frac{\partial^k u_i}{\partial t^k} = \varphi_i^{(k)}(x_1, x_2, \ldots, x_n) \quad (k = 0, 1, 2, \ldots, n_i - 1). \quad (2.2)$$

All the functions $\varphi_i^{(k)}(x_1, x_2, \ldots, x_n)$ are defined on the same region G_0 of the space $(x_1, x_2, \ldots, x_n)$. The function u_i is treated as its own derivative of zero order.

The Cauchy problem consists in finding the solution to the system (2.1) that satisfies the initial conditions (2.2) at $t = t_0$.

The solution is sought in some region G of the space $(t, x_1, \ldots, x_n)$ contained in the region G_0 on the hyperplane $t = t_0$ on which the conditions (2.2) are given.

A particular case of the Cauchy problem is the problem (referred to in the preceding section) of determining the vibrations of an infinite homogeneous membrane from the initial conditions, that is, of finding the solution to the equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2},$$

if, at $t = t_0$

$$u(t_0, x_1, x_2) = \varphi^{(0)}(x_1, x_2) \text{ (initial deviation)},$$

$$u_t'(t_0, x_1, x_2) = \varphi^{(1)}(x_1, x_2) \text{ (initial velocity)}.$$

If $N=1$, $n_1=1$, and $n=0$, the Cauchy problem as formulated earlier is reduced to the problem of finding the solution $u(t)$ of the ordinary differential equation

$$\frac{du}{dt}=F(t, u),$$

that satisfies $u(t_0)=u_0$. This problem is studied in detail in the theory of ordinary differential equations.

2. The function $F(z_1, z_2, \ldots, z_m)$ of m complex variables is said to be analytic in a neighbourhood of the point $z_1^0, z_2^0, \ldots, z_m^0$, if it can be expanded in a power series

$$F(z_1, z_2, \ldots, z_m)= \\ =\sum_{k_1k_2\ldots k_m} A_{k_1k_2\ldots k_m}(z_1-z_1^0)^{k_1}(z_2-z_2^0)^{k_2}\ldots(z_m-z_m^0)^{k_m},$$

that converges for sufficiently small $|z_i-z_i^0|$. It is easy to show that, under these conditions, $F(z_1, z_2, \ldots, z_m)$ has derivatives of all orders at the point $z_1^0, z_2^0, \ldots, z_m^0$ and that

$$A_{k_1k_2\ldots k_m}=\frac{1}{k_1!\,k_2!\ldots k_m!}\left(\frac{\partial^{k_1+k_2+\ldots+k_m}F}{\partial z_1^{k_1}\partial z_2^{k_2}\ldots\partial z_m^{k_m}}\right)_{z_1=z_1^0,\ldots, z_m=z_m^0}.$$

Suppose that $\varphi_i^{(k)}(x_1, x_2, \ldots, x_n)$ are the initial conditions of the Cauchy problem for the system (2.1) (see formula (2.2)). We introduce the abbreviated notations for the derivatives of these functions at a point $(x_1^0, x_2^0, \ldots, x_n^0)$:

$$\left(\frac{\partial^{k-k_0}\varphi_i^{(k_0)}}{\partial x_1^{k_1}\ldots\partial x_n^{k_n}}\right)_{x_1=x_1^0,\ldots, x_n=x_n^0}=\overset{0}{\varphi}_{i,\,k_0,\,k_1,\,k_2,\,\ldots,\,k_n}$$

$$(i=1, 2, \ldots, N;\quad k_0+k_1+\ldots+k_n=k\leqslant n_i).$$

For this we have the following fundamental theorem.

Kovalevsky Theorem. If all the functions F_i are analytic in some neighbourhood of the point $(t_0, x_1^0, \ldots, x_n^0, \ldots \ldots, \overset{0}{\varphi}_{j,k_0,k_1,\ldots,k_n}, \ldots)$ and all the functions $\varphi_j^{(k)}$ are analytic in some neighbourhood of the point $(x_1^0, x_2^0, \ldots, x_n^0)$, then the Cauchy problem has an analytic solution in some neighbourhood of the point $(t^0, x_1^0, x_2^0, \ldots, x_n^0)$ and this solution is unique in the class of analytic functions.

3. We shall prove the Kovalevsky theorem for arbitrary linear systems. The Cauchy problem for such systems is easily reduced to the Cauchy problem for first-order linear systems with the help of a technique that we shall, for simplicity of exposition, illustrate with the example of the single second-order equation

$$\begin{aligned}\frac{\partial^2 u}{\partial t^2} &= \sum_{i,j=1}^{n} a_{ij}(t, x_1, x_2, \ldots, x_n)\frac{\partial^2 u}{\partial x_i \partial x_j} \\ &+ \sum_{i=1}^{n} a_{0i}(t, x_1, \ldots, x_n)\frac{\partial^2 u}{\partial t\, \partial x_i} \\ &+ \sum_{i=1}^{n} b_i(t, x_1, \ldots, x_n)\frac{\partial u}{\partial x_i} + b_0(t, x_1, \ldots, x_n)\frac{\partial u}{\partial t} + \\ &+ c(t, x_1, \ldots, x_n)u + f(t, x_1, \ldots, x_n),\end{aligned} \tag{2.3}$$

where $a_{ij} = a_{ji}$, b_i, c, and f are analytic functions of their arguments in a neighbourhood of the point $(t^0, x_1^0, \ldots, x_n^0)$.

The Cauchy problem for this equation consists in finding the solution satisfying the following initial conditions:

$$\left.\begin{aligned} u(t^0, x_1, \ldots, x_n) &= \varphi_0(x_1, \ldots, x_n), \\ u_t'(t^0, x_1, \ldots, x_n) &= \varphi_1(x_1, \ldots, x_n), \end{aligned}\right\} \tag{2.4}$$

where φ_0 and φ_1 are analytic functions in a neighbourhood of the point $(x_1^0, x_2^0, \ldots, x_n^0)$. We may assume without loss of generality that

$$t^0 = x_1^0 = \ldots = x_n^0 = 0,$$

since the case of arbitrary t^0, $x_1^0, \ldots, x_n^0$ reduces to this case by a change of independent variables that does not change the form of the equation.

If the function $u(t, x_1, \ldots, x_n)$ satisfies equation (2.3) and the initial conditions (2.4), then the functions

$$u,\ u_0 = \frac{\partial u}{\partial t}, \quad u_k = \frac{\partial u}{\partial x_k} \quad (k = 1, 2, \ldots, n)$$

will obviously satisfy the equations

$$\frac{\partial u_0}{\partial t} = \sum_{i,j=1}^{n} a_{ij}\frac{\partial u_i}{\partial x_j} + \sum_{i=1}^{n} a_{0i}\frac{\partial u_0}{\partial x_i} + \sum_{i=1}^{n} b_i u_i + b_0 u_0 + cu + f, \tag{2.5}$$

$$\frac{\partial u_k}{\partial t}=\frac{\partial u_0}{\partial x_k} \quad (k=1,2,\ldots,n), \tag{2.5'}$$

$$\frac{\partial u}{\partial t}=u_0 \tag{2.5"}$$

and the initial conditions

$$u(0, x_1, \ldots, x_n)=\varphi_0(x_1, \ldots, x_n), \tag{2.6}$$

$$u_0(0, x_1, \ldots, x_n)=\varphi_1(x_1, \ldots, x_n), \tag{2.6'}$$

$$u_k(0, x_1, \ldots, x_n)=\frac{\partial \varphi_0(x_1, \ldots, x_n)}{\partial x_k} \tag{2.6"}$$

$$(k=1, \ldots, n).$$

Let us prove the converse: If the functions u, $u_0, u_1, \ldots, u_n$ satisfy equations (2.5), (2.5'), and (2.5") or, more briefly, if they satisfy (2.5) in some region G of the space $(t, x_1, x_2, \ldots, x_n)$ contained in the region G_0 of the space $(x_1, x_2, \ldots, x_n)$ and if they satisfy the initial conditions (2.6), (2.6') and (2.6") in the region G_0, then, throughout the region G, the function $u(t, x_1, x_2, \ldots, x_n)$ satisfies equation (2.3) and the initial conditions (2.4).

It follows from equation (2.5") that throughout the region G,

$$u_0=\frac{\partial u}{\partial t}.$$

If we substitute $\frac{\partial u}{\partial t}$ for u_0 on the right side of equation (2.5') we obtain

$$\frac{\partial u_k}{\partial t}=\frac{\partial^2 u}{\partial t\, \partial x_k} \quad \text{or} \quad \frac{\partial}{\partial t}\left[u_k-\frac{\partial u}{\partial x_k}\right]=0. \tag{2.7}$$

Therefore, the quantity

$$u_k-\frac{\partial u}{\partial x_k}$$

is independent of t throughout the entire region G.

It follows from condition (2.6") that, for $t=0$,

$$u_k=\frac{\partial u}{\partial x_k}$$

in the region G. Therefore it follows* from (2.7) that, for all t,

$$u_k = \frac{\partial u}{\partial x_k} \tag{2.8}$$

throughout the region G. If we make the substitutions $u_0 = \frac{\partial u}{\partial t}$ and $u_k = \frac{\partial u}{\partial x_k}$ in equation (2.5), we see that equation (2.3) is satisfied everywhere in G.

Thus, we have shown that the system (2.5) is equivalent to equation (2.3) if, at $t = 0$,

$$u_k = \frac{\partial u}{\partial x_k}.$$

However, with arbitrary initial conditions, the system (2.5) is, in a certain sense, richer in solutions than is equation (2.3) since arbitrary initial conditions for the solution u, u_0, $u_1, \ldots, u_n$ are not necessarily related by the relations

$$u_k = \frac{\partial u}{\partial x_k}.$$

Problem 1. Show that the Cauchy problem for an arbitrary system (2.1) can be reduced to the Cauchy problem for some first-order system of the form (2.1).

2. Show that the Cauchy problem for a nonlinear first-order system of the form (2.1) can, by differentiation of the equations in the system and introduction of new unknown functions and supplementary equations, be reduced to the Cauchy problem for a quasilinear system of first-order

* Strictly speaking, all that follows from what has been said above is that the difference $u_k - \partial u/\partial r_k$ is independent of t on every segment of a straight line parallel to the t-axis that lies entirely within G. Consequently, $u_k - \partial u/\partial x_k = 0$ in that portion of the region G that is covered by straight line segments parallel to the t-axis lying entirely within G and intersecting G_0. However, since the functions that we are considering are analytic, it follows from a well-known theorem in the theory of analytic functions that these functions vanish identically throughout G.

Obviously, the Cauchy problem for equation (2.3) can be reduced to the Cauchy problem for the system (2.5) by the procedure shown without assuming analyticity of the coefficients of the equation for the initial conditions if the region G is convex with respect to t, that is, if a straight line parallel to the t-axis intersects the boundary of G at not more than two points.

equations, that is, for a system that is linear in all derivatives.

4. Thus, the Cauchy problem for the linear second-order equation (2.3) is reduced to the Cauchy problem for the linear first-order system (2.5). In exactly the same way, we can reduce an arbitrary system of the form (2.1) to a system of first-order equations that are solved explicitly for the derivatives of all the unknown functions with respect to t. Therefore, we shall prove the Kovalevsky theorem for an arbitrary linear system that can be written in the form (2.1) if we can prove it for an arbitrary linear first-order system of the form

$$\frac{\partial u_i}{\partial t}=\sum_{j=1}^{N}\sum_{k=1}^{n} a_{ij}^{(k)}\frac{\partial u_j}{\partial x_k}+\sum_{j=1}^{N} b_{ij}u_j+c_i \qquad (i=1,2,\ldots,N) \tag{2.9}$$

with analytic coefficients under arbitrary analytic initial conditions

$$u_i(0, x_1, \ldots, x_n)=\varphi_i(x_1, x_2, \ldots, x_n) \qquad (i=1,2,\ldots,N). \tag{2.10}$$

The case of arbitrary analytic functions φ_i is easily reduced to the case in which all the functions $\varphi_i(x_1, \ldots, x_n)$ are identically equal to zero. To do this, we introduce new unknown functions

$$v_i(t, x_1, \ldots, x_n)=u_i(t, x_1, \ldots, x_n)-\varphi_i(x_1, \ldots, x_n) \tag{2.11}$$

in place of the original unknown functions $u_i(t, x_1, \ldots, x_n)$. The functions v_i satisfy both the system of equations

$$\frac{\partial v_i}{\partial t}=\sum_{j=1}^{N}\sum_{k=1}^{n} a_{ij}^{(k)}\frac{\partial v_j}{\partial x_k}+\sum_{j=1}^{N} b_{ij}v_j + \left(c_i+\sum_{j=1}^{N}\sum_{k=1}^{n} a_{ij}^{(k)}\frac{\partial \varphi_j}{\partial x_k}+\sum_{j=1}^{N} b_{ij}\varphi_j\right), \tag{2.12}$$

which is completely analogous to the system (2.9), and the initial conditions

$$v_i(0, x_1, x_2, \ldots, x_n)\equiv 0. \tag{2.13}$$

When we prove the existence of a solution to the Cauchy problem for the system (2.12) with zero initial conditions, we shall automatically prove the solvability of the original problem.

For simplicity of writing, we shall assume that the original functions $u_i(t, x_1, \ldots, x_n)$ satisfy the initial conditions

$$u_i(0, x_1, \ldots, x_n) \equiv 0. \tag{2.14}$$

5. Let us first prove the uniqueness of the solution to the Cauchy problem for the system (2.9) with the initial conditions (2.14) in the class of analytic functions in a neighbourhood of the point O with coordinates $t = 0$, $x_1 = 0, \ldots, x_n = 0$. That is, let us show that in no neighbourhood whatever of this point do there exist two analytic solutions of the system (2.9) satisfying the same initial conditions (2.14) at $t = 0$. The functions $u_i(t, x_1, \ldots, x_n)$, which are analytic in a neighbourhood of the coordinate origin, can be expanded close to the origin in series of powers of $t, x_1, \ldots, x_n$. The coefficient $a^i_{k_0 k_1 \ldots k_n}$ of $t^{k_0} x_1^{k_1} \ldots x_n^{k_n}$ in the expansion of the function $u_i(t, x_1, \ldots, x_n)$ is equal to

$$\frac{1}{k_0!\, k_1! \ldots k_n!} \left(\frac{\partial^{k_0 + k_1 + \ldots + k_n} u_i}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}} \right)_{t = x_1 = \ldots = x_n = 0}$$

We shall show that the solution to the Cauchy problem is unique if we can show that the initial conditions (2.14) uniquely determine the coefficients of the expansion (in powers of $t, x_1, \ldots, x_n$) of the functions u_i, that satisfy the system (2.9) or, what amounts to the same thing, if we can show that these conditions uniquely determine the values of all the derivatives of u_i at the point O with coordinates $t = x_1 = \ldots = x_n = 0$. Let us determine these derivatives successively. The initial conditions uniquely determine the values at the point O of all derivatives of the form

$$\left(\frac{\partial^k u_i}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} \right)_{t = x_1 = \ldots = x_n = 0} \tag{2.15}$$

All these derivatives are equal to zero since the identities (2.14) can be differentiated with respect to $x_1, x_2, \ldots, x_n$. Let us assume that there exists a solution to the Cauchy

problem. In equation (2.9), let us replace u_j with the functions constituting this solution. We differentiate all the resulting identities k_1 times with respect to x_1, k_2 times with respect to $x_2, \ldots,$ and k_n times with respect to x_n. Then, on the left sides of these identities, we shall have derivatives of the form

$$\frac{\partial^{k+1} u_i}{\partial t\, \partial x_1^{k_1} \ldots \partial x_n^{k_n}}, \tag{2.16}$$

and, on the right sides, we shall have derivatives with respect to $x_1, x_2, \ldots, x_n$ of the unknown functions and the coefficients of the equations. That is, we shall have quantities that are uniquely determined at the point O by the equations and the initial conditions. The identities obtained determine the values of the derivatives of the form (2.16) (where only one differentiation was with respect to t) at the point O.

Let us differentiate each of the identities (2.9) once with respect to t, k_1 times with respect to $x_1, \ldots,$ and k_n times with respect to x_n. Then, on the right sides of the identities we shall obtain expressions consisting of derivatives of u_i of the form (2.16) and (2.15) and derivatives of the coefficients $a_{ij}^{(k)}$, b_{ij} and c_i. On the left sides, we shall obtain derivatives of the form

$$\frac{\partial^{k+2} u_i}{\partial t^2 \partial x_1^{k_1} \ldots \partial x_n^{k_n}} \tag{2.17}$$

(two differentiations with respect to t). Since we have already shown that derivatives of the form (2.16) and (2.15) are uniquely determined at the point O by the equations (2.9) and the initial conditions (2.14), it follows that all the derivatives (2.17) are uniquely determined at the point O. Repeating this process, we see that all derivatives of u_i are uniquely determined at the point O by equations (2.9) and the initial conditions (2.14). However, the values of all derivatives of the analytic function $u_i(t, x_1, \ldots, x_n)$ at the fixed point O determine uniquely the values of the coefficients in the power series in $t, x_1, \ldots, x_n$ in which this function is expanded about the point O, and hence they completely determine the values of this function in some neighbourhood of the point O. Thus, two analytic solutions of the system (2.9) with the same initial conditions (2.14) must necessarily coincide in some neighbourhood of the coordinate origin.

This proves the uniqueness of the solution to the Cauchy problem for the system (2.9) in the class of analytic functions.

6. In subsection 5, we showed that the initial conditions completely determine the coefficients in the expansion of the functions u_i in powers of $t, x_1, \ldots, x_n$. To prove the existence of a solution to the Cauchy problem, it will be sufficient for us to show that the power series whose coefficients were determined in subsection 5 converge in some neighbourhood of the point O. This is true because if these series converge, the analytic functions $u_i(t, x_1, \ldots, x_n)$ that are defined in terms of them all their partial derivatives with respect to $x_1, x_2, \ldots, x_n$ vanish at the point O (cf. (2.15)). Consequently, they are identically equal to zero at $t = 0$ for all values of $x_1, x_2, \ldots, x_n$, and therefore these functions satisfy the initial conditions (2.14). If we substitute the thus determined functions u_i and their derivatives with respect to $t, x_1, \ldots, x_n$ into equations (2.9), the left sides of these equations will, because of the way in which these functions were constructed at the point O, coincide with the values of the right sides of these equations and their derivatives at that point. Consequently, the left sides of these equations are identically equal to the right sides in some neighbourhood of the coordinate origin, so that these functions satisfy the system (2.9).

To show the convergence of the different powers series that we have obtained for the functions u_i, we will use the method of majorants.

7. A majorant (or majorising function) of a function $\varphi(t, x_1, \ldots, x_n)$ that is analytic in some neighbourhood of a point $(t^0, x_1^0, \ldots, x_n^0)$ is any function $\psi(t, x_1, \ldots, x_n)$ that is analytic in tne same neighbourhood such that all coefficients in the series in powers of $t - t^0, x_1 - x_1^0, \ldots, x_n - x_n^0$ are nonnegative and at least equal in absolute value to the corresponding coefficients in the expansion of the function $\varphi(t, x_1, \ldots, x_n)$.

Let us translate the coordinate origin to the point $(t^0, x_1^0, \ldots, x_n^0)$ and let us construct for the function $\varphi(t, x_1, \ldots, x_n)$ which is assumed analytic in a neighbourhood of the coordinate origin, a majorant of a special form which we shall use in what follows.

coordinate origin, a majorant of a special form which we shall use in what follows.

Suppose that

$$\varphi(t, x_1, \ldots, x_n) = \sum c_{k_0 k_1 \ldots k_n} t^{k_0} x_1^{k_1} \ldots x_n^{k_n}. \tag{2.18}$$

The series on the right converges absolutely at a certain point

$$t = a_0, \; x_1 = a_1, \ldots, x_n = a_n, \quad \text{where} \quad |a_i| > 0$$

for all a_i. Then. there exists a positive number M such that for all non-negative integers $k_0, k_1, \ldots, k_n$

$$|c_{k_0 k_1 \ldots k_n} a_0^{k_0} a_1^{k_1} \ldots a_n^{k_n}| \leqslant M.$$

Consequently, for all $k_0, k_1, \ldots, k_n$

$$|c_{k_0 k_1 \ldots k_n}| \leqslant \frac{M}{|a_0|^{k_0} |a_1|^{k_1} \ldots |a_n|^{k_n}}.$$

Therefore, the function

$$\begin{aligned} S &= \frac{M}{\left(1 - \frac{t}{|a_0|}\right)\left(1 - \frac{x_1}{|a_1|}\right) \ldots \left(1 - \frac{x_n}{|a_n|}\right)} \\ &= M\left[\sum_{k_0=0}^{\infty} \left(\frac{t}{|a_0|}\right)^{k_0} \sum_{k_1=0}^{\infty} \left(\frac{x_1}{|a_1|}\right)^{k_1} \ldots \sum_{k_n=0}^{\infty} \left(\frac{x_n}{|a_n|}\right)^{k_n}\right] \\ &= \sum_{k_0, k_1, \ldots, k_n} \frac{M}{|a_0|^{k_0} |a_1|^{k_1} \ldots |a_n|^{k_n}} t^{k_0} x_1^{k_1} \ldots x_n^{k_n} \end{aligned} \tag{2.19}$$

is a majorant of the function $\varphi(t, x_1, \ldots, x_n)$.

We may employ a different technique for constructing a majorant of a series. For example, the function

$$\frac{M}{1 - \frac{t + x_1 + \ldots + x_n}{a}},$$

where $a = \min(|a_0|, |a_1|, \ldots, |a_n|)$, where $a_i \neq 0$ (for $i = 0, 1, \ldots, n$) and where $(a_0, a_1, \ldots, a_n)$ is some point at which the series (2.18) converges, is also a majorant of the function $\varphi(t, x_1, \ldots, x_n)$ defined by the series (2.18).

To see this, not that, if $|t| + |x_1| + \ldots + |x_n| < a$, this function can be expanded in the series

$$M\sum_{k=0}^{\infty}\left(\frac{t+x_1+\dots+x_n}{a}\right)^k$$
$$=M\sum_{k=0}^{\infty}\frac{1}{a^k}\sum_{k_0+k_1+\dots+k_n=k}\frac{k!}{k_0!\,k_1!\dots k_n!}t^{k_0}x_1^{k_1}\dots x_n^{k_n}; \tag{2.20}$$

But

$$\frac{(k_0+k_1+\dots+k_n)!}{k_0!\,k_1!\dots k_n!}\geqslant 1;\quad \frac{1}{a^k}\geqslant\frac{1}{|a_0|^{k_0}|a_1|^{k_1}\dots|a_n|^{k_n}},$$

that is, the coefficients of our series are positive and at least equal to the corresponding coefficients of the series (2.19). Thus, the function (2.20) is also a majorant for (2.18).

Similarly, the function

$$\frac{M}{1-\dfrac{\frac{t}{\alpha}+x_1+\dots+x_n}{a}}=M\sum_{k=0}^{\infty}\frac{\left(\frac{t}{\alpha}+x_1+\dots+x_n\right)^k}{a^k}, \tag{2.21}$$

where a has its former value and $0<\alpha<1$, is also a majorant of the function $\varphi(t, x_1, \dots, x_n)$.

If we again expand $\left(\frac{t}{\alpha}+x_1+\dots+x_n\right)^k$ in powers of t, $x_1, \dots, x_n$, we shall obtain a series whose coefficients are positive and greater than the corresponding coefficients in the expansion (2.20) in powers of $t, x_1, \dots, x_n$ since the coefficients of the first of these series are obtained from the corresponding coefficients of the second series when we multiply the latter by $\left(\frac{1}{\alpha}\right)^{k_0}$, where $0<\alpha<1$.

Remark 1: Suppose that the power series

$$\varphi(z_1, \dots, z_m)=\sum_{k_1, k_2, \dots, k_m}A_{k_1k_2\dots k_m}z_1^{k_1}z_2^{k_2}\dots z_m^{k_m},$$

converges for $|z_1|\leqslant d_1+\varepsilon, \dots, |z_m|\leqslant d_m+\varepsilon$, where $\varepsilon>0$ is some positive number. Suppose that M^* is the largest absolute value of the function $\varphi(z_1, \dots, z_m)$ as $z_1, \dots, z_m$ assume real and complex values satisfying the conditions

$$|z_1|\leqslant d_1, \dots, |z_m|\leqslant d_m.$$

It can be shown (see V.I. Smirnov. Kurs vysshei matematiki

(Course in Higher Mathematics), Vol. III, Part 2, Section 83, Fizmatgiz, 1958) that the function

$$\frac{M^*}{\left(1-\frac{z_1}{d_1}\right)\left(1-\frac{z_2}{d_2}\right)\dots\left(1-\frac{z_m}{d_m}\right)}$$

majorises the function $\varphi(z_1, \dots, z_m)$. It follows that the function

$$\frac{M^*}{1-\frac{z_1+z_2+\dots+z_m}{d}},$$

where $d = \min(d_1, \dots, d_m)$, also majorises $\varphi(z_1, \dots, z_m)$.

8. We now proceed to prove the existence of a solution to the Cauchy problem for the system (2.9) with initial conditions (2.14). We shall call this 'problem I' and we shall call the system (2.9) 'system I'.

Let us assume that we have majorised the coefficients of the system and the initial conditions of the Cauchy problem. We shall obtain a new system and a new Cauchy problem (which we shall refer to respectively as 'system II' and 'problem II'). Let us show that an analytic solution of problem II will be a majorant of an analytic solution of problem I. If the solution to problem I is represented in a neighbourhood of the origin by the power series

$$u_i = \sum a^{(i)}_{k_0 k_1 \dots k_n} t^{k_0} x_1^{k_1} \dots x_n^{k_n}, \tag{2.22}$$

and the solution to problem II by the series

$$U_i = \sum A^{(i)}_{k_0 k_1 \dots k_n} t^{k_0} x_1^{k_1} \dots x_n^{k_n}, \tag{2.23}$$

we need to prove the inequality between the coefficients

$$\left|a^{(i)}_{k_0 k_1 \dots k_n}\right| \leqslant A^{(i)}_{k_0 k_1 \dots k_n}. \tag{2.24}$$

In the case of $k_0 = 0$, these inequalities follow immediately from the fact that the initial conditions of problem II majorise the initial conditions of problem I. For the case in which $k_0 > 0$, the coefficients $a^{(i)}_{k_0 k_1 \dots k_n}$, (resp. $A^{(i)}_{k_0 k_1 \dots k_n}$,)

are obtained from the coefficients $a^{(i)}$ (resp. $A^{(i)}$) that have smaller subscript k_0 and from the values of the coefficients in the system I (resp. II) and their derivatives at the point O. Therefore, it is easy to see that if inequalities (2.24) hold for $k_0 < k$, they hold also for $k_0 = k$. This means that they hold for all coefficients in the expansions (2.22) and (2.23).

Consequently, solvability of problem II (convergence of the series (2.22)) implies solvability of problem I (convergence of the series (2.23)). However, problem II can be constructed with a great degree of arbitrariness since we can arbitrarily choose majorants for the coefficients and initial conditions of problem I. Let us choose problem II sufficiently simple for use to be able to find its solution easily. We take numbers $M > 0$ and $a > 0$ such that the functions

$$\frac{M}{1 - \dfrac{\dfrac{t}{\alpha} + x_1 + \ldots + x_n}{a}}$$

where $0 < \alpha < 1$ will majorise all coefficients of the system except the free terms. For the free terms, we take a general majorant of the form*

$$\frac{M_1}{1 - \dfrac{\dfrac{t}{\alpha} + x_1 + \ldots + x_n}{a}}$$

This is possible since a majorant of this form exists for every coefficient and, to construct a general majorant, we take M and M_1 equal to the largest and a equal to the smallest of all their values corresponding to the different coefficients. Now that we have chosen the numbers M, M_1, and a as described, we write the majorising system in the form

$$\frac{\partial U_i}{\partial t} = \frac{M}{1 - \dfrac{\dfrac{t}{\alpha} + x_1 + \ldots + x_n}{a}} \left[\sum_{j=1}^{N} \sum_{k=1}^{n} \frac{\partial U_j}{\partial x_k} + \sum_{j=1}^{N} U_j + m \right] \quad (2.25)$$

where the number α (such that $0 < \alpha < 1$) will be chosen

* The fact that we can choose M_1 independently of M will be very useful to us in what follows (see remark 2 at the end of this section).

later and $m=\frac{M_1}{M}$.

Without at the moment fixing the initial conditions, let us seek a solution of the system in the form

$$\begin{aligned} U_1(t, x_1, \ldots, x_n) &\equiv U_2(t, x_1, \ldots, x_n) \equiv \ldots \\ \ldots &\equiv U_N(t, x_1, \ldots, x_n) = U(t, x_1, \ldots, x_n) \\ &= U\left(\frac{t}{a}+x_1+\ldots+x_n\right) = U(z), \end{aligned}$$

where

$$z=\frac{t}{a}+x_1+\ldots+x_n.$$

If we substitute the postulated solution into the system (2.25), we see that the function $U(z)$ must satisfy the equation

$$\frac{1}{a}\frac{dU}{dz}=A(z)\left(Nn\frac{dU}{dz}+NU+m\right), \tag{2.26}$$

where

$$A(z)=\frac{M}{1-\frac{z}{a}}.$$

The variables in this equation are separable, so that we may write

$$\frac{dU}{\frac{N}{m}U+1}=\frac{mA(z)\,dz}{\frac{1}{a}-NnA(z)}=B(z)\,dz.$$

We now choose a positive number a sufficiently small that, in some neighbourhood of the point $z=0$

$$\frac{1}{a}-NnA(z)>0. \tag{2.27}$$

Then, $B(z)$ will be an analytic function in this neighbourhood.

Let us show that the particular solution of equation (2.26)

$$U(z)=\frac{e^{\frac{N}{m}\int_0^z B(\xi)\,d\xi}-1}{N}\,m$$

yields the desired majorant for the solution of problem I.

Since the function

$$U_i(t, x_1, \ldots, x_n)=U\left(\frac{t}{a}+x_1+\ldots+x_n\right)$$

satisfy the system (2.25), which majorises the original system, all that we need to do to prove this assertion is to show that $U(z)$ can be expanded in a series in $x_1, x_2, \ldots, x_n$ with positive coefficients at $t=0$ that is, that it majorises the identically-zero function (the initial conditions of problem I).

To show this, note that

$$A(z)=\frac{M}{1-\frac{z}{a}}$$

is a function whose expansion in powers of z has non-negative coefficients. Consequently,

$$B(z)=\frac{m\alpha A(z)}{1-\alpha A(z)Nn}$$
$$=m\alpha A(z)\left[1+\alpha NnA(z)+\alpha^2N^2n^2A^2(z)+\ldots\right]$$

also has non-negative coefficients when it is expanded in powers of z. Therefore,

$$C(z)=\frac{N}{m}\int_0^z B(z)\,dz,\quad e^{C(z)}-1=C(z)+\frac{C^2(z)}{2!}+\ldots,\quad U(z)$$

also possesses this property. Therefore, the coefficients of the expansion of $U(x_1+x_2+\ldots+x_n)$ in powers of $x_1, x_2, \ldots, x_n$ are also non-negative; that is, $U(0, x_1, x_2, \ldots, x_n)$ is indeed a majorant of zero. This means that the functions

$$U_i(t, x_1, \ldots, x_n)=U\left(\frac{t}{\alpha}+x_1+\ldots+x_n\right)$$

are the solution of some problem II. The analyticity of this solution follows from the fact that, as was shown above, $U(z)$ can be expanded in a series of powers of z and, consequently, in a series of powers of $t, x_1, \ldots, x_n$. And from this, as was shown above, follows the convergence of the power series (2.22) that represent the solution of the original problem. This completes the proof of the Kovalevsky theorem for linear systems.

Remark 2: It is clear from the proofs of the theorem that the series representing the solution to the Cauchy problem for the system (2.9) always converge in the region in which the series representing the solution of the majorising problem converge. It follows from this that the solution of the

original Cauchy problem for the system (2.9) and the original functions φ_i, not necessarily equal to zero, always exists in some region

$$\left|\frac{t}{\alpha}\right| < \rho, \ |x_1| < \rho, \ \ldots, \ |x_n| < \rho, \ \rho > 0,$$

if the coefficients of the system (2.9) and the original functions are holomorphic in the region

$$|t| \leqslant R, \ |x_i| \leqslant R \qquad (i = 1, 2, \ldots, n, \ R > 0).$$

Here, ρ and α depend only on R and the number M; they do not depend at all on the values of the original functions φ_i or the free terms of the equations since neither α nor the region of variation of z in which inequality (2.27) holds depends on these values.

Remark 3: The Kovalevsky theorem is not generally applicable for systems that are not of the form (2.1), as can be seen from the following example, exhibited by Madame Kovalevsky herself. Consider the equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{2.28}$$

with initial conditions

$$u(0, x) = \frac{1}{1-x}, \qquad |x| < 1. \tag{2.29}$$

It is easy to see that if an analytic solution $u(t, x)$ to the problem (2.28)-(2.29) exists, it must be representable in a neighbourhood of the coordinate origin by the series

$$\sum_{n=0}^{\infty} \frac{(2n)!}{n!} \frac{t^n}{(1-x)^{2n+1}};$$

However, this series diverges at every point such that $t \neq 0$.

Problem. Prove the Kovalevsky theorem for a quasilinear system of first-order equations.

3. GENERALISATION OF THE CAUCHY PROBLEM.

1. Consider a system of N equations with N unknown functions $u_1, u_2, \ldots, u_N$

For each of these functions u_i, there is an integer n_i; re-

$$\Phi_i\left(x_0, x_1, \ldots, x_n; u_1, \ldots, u_N, \ldots, \frac{\partial^k u_j}{\partial x_0^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}}, \ldots\right)=0 \qquad (3.1)$$
$$(i, j=1, 2, \ldots, N).$$

presenting the highest-order partial derivative of u_i with respect to the independent variables $x_0, x_1, \ldots, x_n$, that appears in the system (3.1). Suppose that a sufficiently smooth n-dimensional surface S is given in the region in question of the points $(x_0, x_1, \ldots, x_n)$ and that at each point of the surface there is some line l, not tangent to S, that changes sufficiently smoothly from point to point along S, for example, the normal to the surface. Suppose that all the functions u_i and their derivatives in the direction of the line l up to order $n_i - 1$ are given on that surface. These conditions on the surface S constitute a generalisation of the Cauchy (initial) conditions that we examined in the preceding section. We are required to find a solution $u_1, u_2, \ldots, u_N$ of the system (3.1) in some neighbourhood of the surface that will satisfy the conditions given on S.

2. Let us try to reduce this problem to the Cauchy problem formulated in the preceding section. For simplicity, we shall at first confine ourselves to examining not the system (3.1) but the following linear system:

$$\sum_{j, k_0, k_1, \ldots, k_n} A_{ij}^{(k_0 \ldots k_n)}(x_0, x_1, \ldots, x_n)\frac{\partial^{n_j} u_j}{\partial x_0^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}} + \qquad (3.2)$$
$$\ldots + f_i(x_0, x_1, \ldots, x_n) = 0$$
$$(i, j = 1, 2, \ldots, N).$$

Here we have written only the terms with highest-order derivatives of the unknown functions.

In a neighbourhood of the surface S, we introduce new curvilinear coordinates $\xi_0, \xi_1, \ldots, \xi_n$ in such a way that the equation of the surface S will take the form $\xi_0 = 0$ and the lines l will coincide with the coordinate curves

$$\xi_1 = c_1, \ \xi_2 = c_2, \ \ldots, \ \xi_n = c_n.$$

To do this, we shall make more precise the assumptions made regarding the smoothness of the surface S and the lines l. We assume that on the surface we may introduce the

parameters $\xi_1, \ldots, \xi_n$:

$$\left.\begin{aligned} x_0 &= x_0(\xi_1, \ldots, \xi_n), \\ x_1 &= x_1(\xi_1, \ldots, \xi_n), \\ &\ldots\ldots\ldots\ldots \\ x_n &= x_n(\xi_1, \ldots, \xi_n), \end{aligned}\right\} \tag{3.3}$$

such that the rank of the Jacobian matrix

$$\left\|\frac{\partial x_i}{\partial \xi_k}\right\| \quad (i = 0, 1, \ldots, n;\ k = 1, 2, \ldots, n)$$

will be equal to n at every point of S and such that the right sides of equations (3.3) will be sufficiently smooth functions.

Suppose that the lines l are given by the parametric equations

$$\left.\begin{aligned} x_0 &= X_0(\xi_0, \xi_1, \ldots, \xi_n), \\ &\ldots\ldots\ldots\ldots\ldots \\ x_n &= X_n(\xi_0, \xi_1, \ldots, \xi_n), \end{aligned}\right\} \tag{3.4}$$

where ξ_0 is the parameter of the point along the line l and $\xi_1, \ldots, \xi_n$ are the parameters of the point of intersection of l with S. Then, the right sides of equations (3.4) are assumed to be sufficiently smooth functions of all their arguments.

With regard to the parameter ξ_0, we assume that at least one of the derivatives $\frac{\partial X_i}{\partial \xi_0}$ (for $i = 0, 1, \ldots, n$) is nonzero and that the value $\xi_0 = 0$ corresponds to the point of intersection of the line l with the surface S (that is, that equations (3.4) coincide for $\xi_0 = 0$ with equations (3.2) of the surface S).

Let us now show that the Jacobian determinant

$$I = \begin{vmatrix} \frac{\partial X_0}{\partial \xi_0} & \frac{\partial X_1}{\partial \xi_0} & \cdots & \frac{\partial X_n}{\partial \xi_0} \\ \frac{\partial X_0}{\partial \xi_1} & \frac{\partial X_1}{\partial \xi_1} & \cdots & \frac{\partial X_n}{\partial \xi_1} \\ \cdot & \cdot & \cdots & \cdot \\ \frac{\partial X_0}{\partial \xi_n} & \frac{\partial X_1}{\partial \xi_n} & \cdots & \frac{\partial X_n}{\partial \xi_n} \end{vmatrix} \tag{3.5}$$

is nonzero in some neighbourhood of the surface S. On the

surface S, that is, for $\xi_0 = 0$, we have

$$I = \begin{vmatrix} \frac{\partial X_0}{\partial \xi_0} & \frac{\partial X_1}{\partial \xi_0} & \cdots & \frac{\partial X_n}{\partial \xi_0} \\ \frac{\partial x_0}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_1} & \cdots & \frac{\partial x_n}{\partial \xi_1} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial x_0}{\partial \xi_n} & \frac{\partial x_1}{\partial \xi_n} & \cdots & \frac{\partial x_n}{\partial \xi_n} \end{vmatrix}. \tag{3.6}$$

The last n rows of the determinant (3.6) are linearly independent since, by hypothesis, the rank of the matrix

$$\left\| \frac{\partial x_i}{\partial \xi_k} \right\| \quad (i = 0, 1, \ldots, n; \ k = 1, \ldots, n)$$

is equal to n. If the determinant (3.6) were equal to zero, its first row, representing a nonzero vector tangential to l, would be a linear combination of the remaining n rows. But this is impossible since these last n rows are vectors lying on the hyperplane tangent to S and the lines l are assumed not to be tangent to S.

Because of continuity, the determinant (3.5) is nonzero in some neighbourhood of S. Therefore, in this neighbourhood we may take $\xi_0, \xi_1, \ldots, \xi_n$ as our new coordinates of the point $(x_0, x_1, \ldots, x_n)$.

Let us now turn to the independent variables $\xi_0, \xi_1, \ldots, \xi_n$ in equations (3.2). In the transformed equations, we shall be especially interested in the terms containing n_ith-order derivatives of the u_i with respect to ξ_0 (i.e. the highest-order derivatives). If we write only these terms, we obtain

$$\frac{\partial^{n_i} u_i}{\partial x_0^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}} = \frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}} \left(\frac{\partial \xi_0}{\partial x_0} \right)^{k_0} \left(\frac{\partial \xi_0}{\partial x_1} \right)^{k_1} \cdots \left(\frac{\partial \xi_0}{\partial x_n} \right)^{k_n} + \ldots$$

Therefore, if we write only the terms containing the highest-order derivatives of the functions u_i with respect to ξ_0 in the equations obtained from the transformation of equations (3.2), we obtain

$$\sum_{\substack{j=1 \\ k_0 + \ldots + k_n = n_j}}^{N} A_{ij}^{(k_0 \ldots k_n)} \left(\frac{\partial \xi_0}{\partial x_0} \right)^{k_0} \cdots \left(\frac{\partial \xi_0}{\partial x_n} \right)^{k_n} \frac{\partial^{n_j} u_j}{\partial \xi_0^{n_j}} + \ldots = 0 \tag{3.7}$$

$$(i = 1, 2, \ldots, N).$$

For these equations to have a unique solution (for the derivatives $\frac{\partial^{n_j} u_j}{\partial \xi_0^{n_j}}$) close to the surface S no matter what the remaining terms of the equation (which we have not written out) maybe, it is necessary and sufficient that the determinant

$$\left| \sum_{k_0+\ldots+k_n=n_j} A_{ij}^{(k_0 k_1 \ldots k_n)} \left(\frac{\partial \xi_0}{\partial x_0}\right)^{k_0} \cdots \left(\frac{\partial \xi_0}{\partial x_n}\right)^{k_n} \right|$$
$$(i, j=1, 2, \ldots, N).$$

be nonzero at all points of the surface S. Then, because of the continuity of the coefficients $A_{ij}^{(k_0 \ldots k_n)}$ and of the derivatives $\frac{\partial \xi_0}{\partial x_k}$, this determinant will also be nonzero in some neighbourhood of the surface S in the space $(x_0, x_1, \ldots, x_n)$.

The equation

$$\left| \sum_{k_0+\ldots+k_n=n_j} A_{ij}^{(k_0 k_1 \ldots k_n)} (x_0, \ldots, x_n)\, \alpha_0^{k_0} \alpha_1^{k_1} \ldots \alpha_n^{k_n} \right| = 0 \qquad (3.8)$$
$$(i, j=1, 2, \ldots, N)$$

is called the characteristic equation for the system (3.2). Here, $\alpha_0, \alpha_1, \ldots, \alpha_n$ are parameters and $\sum_{i=0}^{n} \alpha_i^2 \neq 0$. The direction of the hyperplane

$$\sum_{k=0}^{n} \alpha_k (x_k - x_k^0) = 0$$

is called the characteristic direction at the point $(x_0^0, \ldots, x_n^0)$ for the system (3.2) if*

* Since equation (3.8) is homogeneous with respect to the unknowns $\alpha_0, \alpha_1, \ldots, \alpha_n$, these unknowns can be normalised if we take, for example,

$$\sum_{k=0}^{n} \alpha_k^2 = 1.$$

Then α_k will be the cosine of the angle between the normal to the characteristic hyperplane and the x_k-axis.

$$\left| \sum_{k_0+\ldots+k_n=n_j} A_{ij}^{(k_0 \ldots k_n)} (x_0^0, x_1^0, \ldots, x_n^0) \alpha_0^{k_0} \ldots \alpha_n^{k_n} \right| = 0$$

The surface $\varphi(x_0, x_1, \ldots, x_n) = 0$ is called the characteristic surface for the system (3.2) or simply the characteristic if at every point of this surface

$$\left| \sum_{k_0+\ldots+k_n=n_j} A_{ij}^{(k_0 k_1 \ldots k_n)} (x_0, x_1, \ldots, x_n) \times \left(\frac{\partial \varphi}{\partial x_0}\right)^{k_0} \left(\frac{\partial \varphi}{\partial x_1}\right)^{k_1} \ldots \left(\frac{\partial \varphi}{\partial x_n}\right)^{k_n} \right| = 0$$

and if at least one of the derivatives $\frac{\partial \varphi}{\partial x_k}$ $(k = 0, 1, \ldots, n)$ is nonzero.

It follows from these definitions that the direction of every tangent hyperplane to the characteristic surface, or, as we shall say for brevity, the direction of the characteristic surface, is everywhere characteristic.

3. It is clear from what was said above that if the direction of the surface S of which we were speaking in the formulation of the generalised Cauchy problem is nowhere characteristic for the system (3.2), then, after the coordinates $\xi_0, \xi_1, \ldots, \xi_n$ are introduced in place of $x_0, x_1, \ldots, x_n$, as described in Section 2, the transformed system (3.7) can always be solved for the highest-order derivatives of the u_i with respect to ξ_0 close to the surface S. We obtain the system

$$\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}} = \sum_{j,k} B_{ij}^{(k_0 \ldots k_n)} (\xi_0, \ldots, \xi_n) \frac{\partial^k u_j}{\partial \xi_0^{k_0} \ldots \partial \xi_n^{k_n}} + F_i(\xi_0, \ldots, \xi_n) \tag{3.9}$$

$$(i, j = 1, 2, \ldots, N;\ k = k_0 + k_1 + \ldots + k_n \leqslant n_j;\ k_0 < n_j).$$

The conditions given on the surface S then become the conditions

$$\left(\frac{\partial^k u_i}{\partial \xi_0^k}\right)_{\xi_0=0} = \varphi_i^{(k)}(\xi_1, \ldots, \xi_n); \quad \begin{array}{l} k = 0, 1, \ldots, n_i - 1, \\ i = 1, 2, \ldots, N. \end{array} \tag{3.10}$$

Thus, if the surface S had no characteristic direction anywhere, the generalised Cauchy problem would reduce to the

original Cauchy problem. The change from the first of these problems to the second is completely reversible: to every sufficiently smooth solution* of one problem there corresponds a unique smooth solution of the other.

However, in the preceding section, we were speaking of the solution of a system with analytic coefficients and analytic initial conditions. For the system (3.9) and its Cauchy problem to satisfy these requirements, it is sufficient that the following supplementary conditions be satisfied: (a) The coefficients of the system (3.2) are analytic functions of $x_0, \ldots, x_n$. (b) The functions $x_i = X_i(\xi_0, \xi_1, \ldots, \xi_n)$, for $i = 0, 1, \ldots, n$, are analytic functions of their arguments. The possibility of choosing analytic functions X_i is connected with the nature of the surface S and of the family of lines l. We shall call a surface S and a family of lines l for which this is possible an analytic surface and an analytic family of lines.

If the surface is defined by the equation $F(x_0, x_1, \ldots, x_n) = 0$, it will be analytic in the case in which the function $F(x_0, x_1, \ldots, x_n)$ is an analytic function of its arguments and the surface does not contain singular points (that is, points at which all the first-order derivatives of the function F vanish). The family of normals to an analytic surface is an analytic family of lines. (c) The initial conditions are analytic functions of $\xi_1, \ldots, \xi_n$.

According to the Kovalevsky theorem, we can assert that when conditions (a), (b), and (c) are satisfied, the generalised Cauchy problem always has a unique solution in some neighbourhood of the surface S if this surface has no characteristic direction anywhere.

On the other hand, if the surface S has a characteristic direction at some point A, that is, if the equation

$$\left| \sum_{k_0+\ldots+k_n=n_j} A_{ij}^{(k_0 \ldots k_n)} \left(\frac{\partial \xi_0}{\partial x_0}\right)^{k_0} \cdots \left(\frac{\partial \xi_0}{\partial x_n}\right)^{k_n} \right| = 0, \tag{3.11}$$

holds at a point A on the surface

$$\xi_0(x_0, \ldots, x_n) = 0$$

* It is sufficient to require that the functions u_i have continuous derivatives of order n_i and that the functions defining the transformation of coordinates have continuous derivatives of order $\max(n_i)$.

then we cannot in general give arbitrary values to the functions u_i and their derivatives on the surface S and still be sure that the generalised Cauchy problem has a solution. To see this, let us leave on the left sides of equations (3.7) all terms containing derivatives of order n_i of the function u_i with respect to ξ_0 and let us transpose all the remaining terms to the right. By virtue of condition (3.11) at the point A, there will then be a linear dependence between the left sides of the equations so formed. This means that the same linear dependence must exist between the right sides of these equations, which are completely determined by the given values of the functions u_i and their derivatives on the surface S. But this imposes a definite dependence on these initial conditions provided the required linear dependence between the right sides of the equations is not identically satisfied for all values of the functions u_i and their derivatives on S. In the latter case and also in the case in which the conditions of the Cauchy problem given on the characteristic surface S are such that the system has a solution for the highest-order derivatives with respect to ξ_0, which are treated as functions of the unknown variables

$$\xi_0, \ldots, \xi_n, u_1, \ldots, u_N, \frac{\partial^k u_j}{\partial \xi_0^{k_0} \ldots \partial \xi_n^{k_n}}$$
$$(k = \sum k_s \leqslant n_j, \ k_0 < n_j, \ j = 1, \ldots, N),$$

such a solution may not be unique in a neighbourhood of the point A.

We shall give some examples for finding the characteristic directions for the equations and systems of equations. Here, we shall always assume that

$$\sum \alpha_i^2 = 1, \tag{3.12}$$

that is, that α_i denotes the cosine of the angle between the x_i-axis and the normal to the hyperplane having the characteristic direction.

Example 1. For Laplace's equation

$$\frac{\partial^2 u}{\partial x_0^2} + \ldots + \frac{\partial^2 u}{\partial x_n^2} = 0$$

equation (3.8) takes the form

$$\alpha_0^2 + \alpha_1^2 + \ldots + \alpha_n^2 = 0.$$

Keeping equation (3.12) in mind, we see that Laplace's equation does not have real characteristic directions.

Example 2. For the wave equation

$$\frac{\partial^2 u}{\partial x_0^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$$

equation (3.8) takes the form

$$\alpha_0^2 - \alpha_1^2 - \alpha_2^2 = 0.$$

Since by virtue of (3.12)

$$\alpha_0^2 + \alpha_1^2 + \alpha_2^2 = 1,$$

it follows that $2\alpha_0^2 = 1$, so that $\alpha_0 = \pm\frac{1}{\sqrt{2}}$. This means that the tangent planes to all characteristic surfaces make a 45° angle with the x_0-axis. By using this property of the characteristic surfaces, we easily see what form the characteristic surfaces passing through certain curves on the $(x_0 = \text{const})$-plane must have. For example, the plane passing through l and forming a 45° angle with the $(x_0 = \text{const})$-plane is a characteristic surface passing through an arbitrary straight line l lying in that plane. The lateral surface of a circular cone with axis parallel to the x_0-axis and with generators making a 45° angle with the $(x_0 = \text{const})$-plane or, what amounts to the same thing, with the x_0-axis is a characteristic surface passing through any circle K in the $(x_0 = \text{const})$-plane.

It is easy to see that completely analogous results hold for the so-called wave equation in n-dimensional space:

$$\frac{\partial^2 u}{\partial x_0^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \ldots + \frac{\partial^2 u}{\partial x_n^2}$$

3. For the heat-flow equation

$$\frac{\partial u}{\partial x_0} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \ldots + \frac{\partial^2 u}{\partial x_n^2}$$

equation (3.8) takes the form

$$\alpha_1^2 + \ldots + \alpha_n^2 = 0.$$

On the basis of (3.12), it follows from this that $\alpha_0^2 = 1$.

Therefore, the hyperplanes $x_0 = \text{const}$ are characteristic surfaces.

4. For the equation

$$a_1(x_1, \ldots, x_n)\frac{\partial u}{\partial x_1} + a_2(x_1, \ldots, x_n)\frac{\partial u}{\partial x_2} \ldots + a_n(x_1, \ldots, x_n)\frac{\partial u}{\partial x_n} = 0$$

the relation (3.8) takes the form

$$a_1(x_1, \ldots, x_n)\alpha_1 + a_2(x_1, \ldots, x_n)\alpha_2 \ldots + a_n(x_1, \ldots, x_n)\alpha_n = 0.$$

Therefore, all hyperplanes passing through the point $(x_1, \ldots, x_n)$ and through the vector originating at this point with components $(a_1, \ldots, a_n)$ have a characteristic direction at that point.

Example 5. For a system of equations with two independent variables

$$\sum_{j=1}^{n} a_{ij}(x_1, x_2)\frac{\partial u_j}{\partial x_1} + \sum_{j=1}^{n} b_{ij}(x_1, x_2)\frac{\partial u_j}{\partial x_2} + \sum_{j=1}^{n} c_{ij}(x_1, x_2)u_j = 0$$
$$(i = 1, 2, \ldots, n)$$

equation (3.8) takes the form

$$|\alpha_1 a_{ij}(x_1, x_2) + \alpha_2 b_{ij}(x_1, x_2)| = 0.$$

The curves along which $\frac{dx_2}{dx_1}$, that is, $-\frac{\partial\varphi}{\partial x_1}\Big/\frac{\partial\varphi}{\partial x_2}$, is equal to some root k of the equation

$$|-ka_{ij}(x_1, x_2) + b_{ij}(x_1, x_2)| = 0$$

will be characteristic curves.

Here, we assume that $\varphi(x_1, x_2) = \text{const}$ is the equation of the characteristic line.

Problem. Show that for a smooth nondegenerate transformation of coordinates the characteristic surface of the system (3.2) becomes a characteristic surface of the transformed system; that is, show that the characteristics are invariant under nondegenerate transformation of coordinates.

4. For nonlinear systems of the form

$$\Phi_i\left(x_0, \ldots, x_n;\ u_1, \ldots, u_N, \ldots, \frac{\partial^k u_j}{\partial x_0^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}}, \ldots\right)=0 \tag{3.13}$$

$$(i, j=1, \ldots, N;\ k=k_0+k_1+\ldots+k_n \leqslant n_j)$$

the equation

$$\left| \sum_{k_0+k_1+\ldots+k_n=n_j} \frac{\partial \Phi_i}{\partial \left\{ \dfrac{\partial^{n_j} u_j}{\partial x_0^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}} \right\}} \alpha_0^{k_0} \ldots \alpha_n^{k_n} \right| = 0 \tag{3.14}$$

is called the characteristic equation. The surface

$$\varphi(x_0, \ldots, x_n)=0 \tag{3.15}$$

is said to be characteristic for the system (3.13) and for the given solution $u_1, u_2, \ldots, u_N$ of this system if the following identity holds on this surface for the functions $u_1, u_2, \ldots, u_N$:

$$\left| \sum_{k_0+k_1+\ldots+k_n=n_j} \frac{\partial \Phi_i}{\partial \left\{ \dfrac{\partial^{n_j} u_j}{\partial x_0^{k_0} \ldots \partial x_n^{k_n}} \right\}} \left(\frac{\partial \varphi}{\partial x_0}\right)^{k_0} \ldots \left(\frac{\partial \varphi}{\partial x_n}\right)^{k_n} \right| = 0.$$

The characteristic equations for the system (3.13) are defined analogously at the point in space $(x_0, \ldots, x_n)$ for the given solution $u_1, u_2, \ldots, u_N$. In the case of nonlinear systems, it is meaningful to speak of the characteristic direction of the hyperplane

$$\sum \alpha_k (x_k - \overset{0}{x}_k) = 0$$

at the given point only for a definite solution $u_1, u_2, \ldots, u_N$ of the system (3.13) since the coefficients of equation (3.14) generally depend in this case on the functions u_i and their first n_i; derivatives.

By analogy with what we did in subsection 3, we may say the following: Suppose that the Cauchy conditions are given on some analytic surface as was done in the preceding case and suppose that all the functions involved are analytic. Since the system in question is not now assumed to be linear, after changing to the coordinates $\xi_0, \xi_1, \ldots, \xi_n$, as was done

in the preceding case, we obtain for $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$ a nonlinear system of equations. Let us denote it by $\sum$. In general, this system does not have a unique solution for the $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$, (where $i = 1, 2, \ldots, N$), which are treated as functions of of the independent variables

$$\xi_0, \ldots, \xi_n, \ldots, u_j, \ldots, \frac{\partial^k u_j}{\partial \xi_0^{k_0} \ldots \partial \xi_n^{k_n}},$$

where

$$k = \sum k_s \leqslant n_j, \quad k_0 < n_j$$

(for $j = 1, 2, \ldots, N$). Let us assume that close to the hypersurface $\xi_0 = 0$ and approximate to the values of u_j and their derivatives that are given on it, we have chosen for the $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$ (where $i = 1, 2, \ldots, N$) some system of analytic functions of

$$\xi_0, \ldots, \xi_n, \ldots, u_j, \ldots, \frac{\partial^k u_j}{\partial \xi_0^{k_0} \ldots \partial \xi_n^{k_n}}, k = \sum k_s \leqslant n_j, k_0 < n_j$$

(for $j = 1, 2, \ldots, N$) satisfying equations $\sum$. Thus, we determine the values $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$ on the surface S from the initial conditions of the generalised Cauchy problem that are given on it. If we now return to the coordinates $x_0, \ldots, x_n$, we obtain the values of all the functions u_i and their first n_i derivatives with respect to $x_0, \ldots, x_n$ on the surface S. If we replace the u_i and their derivatives in equation (3.14), with these values, we obtain a completely determined equation for $\alpha_0, \alpha_1, \ldots, \alpha_n$. Consequently, we may thus determine the characteristic directions at every point $(x_0, x_1, \ldots, x_n)$ of the surface S. Let us suppose that the surface S has no characteristic direction anywhere. Then, we may show that the generalised Cauchy problem that is thus posed for the system (3.13) has a unique analytic solution for the choice made for the values of the $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$ on S.

4. THE UNIQUENESS OF THE SOLUTION TO THE CAUCHY PROBLEM IN THE CLASS OF NONANALYTIC FUNCTIONS

1. The existence and uniqueness of the solution to the Cauchy problem in the class of analytic functions follow from the Kovalevsky theorem if analytic Cauchy conditions are given on an analytic surface S that nowhere has a characteristic direction. It follows from the constructions made in Sections 2 and 3 that if all functions appearing in the given equations and in the initial conditions assume real values for real values of the arguments, the solutions to the Cauchy problem will also be real. The question then arises: in this case, are there not other solutions to the Cauchy problem besides Kovalevsky's analytic solution? Note that for a system of functions $(u_1, \ldots, u_N)$ to be a real solution to the Cauchy problem, it is not necessary to require that all the functions u_i be analytic. For this, it is quite sufficient for them to possess derivatives of all orders that appear in the equations in question. Despite the efforts of many outstanding mathematicians, this question has not as yet been completely solved.

As far back as 1901, Holmgren showed that the solution of the Cauchy problem with initial conditions (3.10) for linear systems of equations of the form (3.9) with analytic coefficient is unique in the class of functions possessing continuous derivatives of all orders that appear in the system in question.

We shall now prove this theorem. For simplicity in exposition, we shall assume that the number of independent variables is two although essentially the same proof is applicable for an arbitrary number of independent variables. We also assume that the system in question is of first order. From what was said in Section 2, the general case can be reduced to this one. We denote the independent variables by $\breve{x}$ and y and we assume first that the Cauchy problem is stated for an interval of the straight line $\breve{x}=0$ that contains the coordinate origin.

Thus, suppose that we are given the system of equations

$$\frac{\partial z_i}{\partial \breve{x}} = \sum_{j=1}^{n} A_{ij}(\breve{x}, y)\frac{\partial z_j}{\partial y} + \sum_{j=1}^{n} B_{ij}(\breve{x}, y) z_j + C_i(\breve{x}, y) \qquad (i = 1, 2, \ldots, n) \tag{4.1}$$

and the initial conditions

$$z_i(0, y) = \varphi_i(y). \tag{4.2}$$

A_{ij}, B_{ij}, and C_i are analytic functions of their arguments in some neighbourhood of the coordinate origin. Suppose that close to the coordinate origin two solutions of the system (4.1) are given that satisfy these same initial conditions (4.2) and that these solutions consist of the functions $z_1, \ldots, z_n$ (first solution) and $\breve{z}_1, \ldots, \breve{z}_n$ (second solution) possessing continuous first-order partial derivatives. We need to show that these solutions coincide in some neighbourhood of the origin.

Let us set

$$z_i - \breve{z}_i = \breve{u}_i \quad (i = 1, 2, \ldots, n).$$

Then, all the $\breve{u}_i$ are continuously differentiable functions close to the origin and they satisfy the equations

$$\frac{\partial \breve{u}_i}{\partial \breve{x}} = \sum_{j=1}^{n} A_{ij}(\breve{x}, y)\frac{\partial \breve{u}_j}{\partial y} + \sum_{j=1}^{n} B_{ij}(\breve{x}, y)\,\breve{u}_j \quad (i = 1, 2, \ldots, n)$$

and the initial conditions

$$\breve{u}_i(0, y) = 0 \qquad (i = 1, 2, \ldots, n).$$

Let us show that all the $\breve{u}_i$ are identically equal to zero close to the point O (0, 0). Instead of $\breve{x}$ we introduce a new independent variable

$$x = \breve{x} + y^2$$

and we set

$$u_i(x, y) = \breve{u}_i(\breve{x}, y) = \breve{u}_i(x - y^2, y) \qquad (i = 1, 2, \ldots, n).$$

Then, the functions u_i satisfy the system of equations

$$\frac{\partial u_i}{\partial x} = 2y\sum_{j=1}^{n} A_{ij}(x - y^2, y)\frac{\partial u_j}{\partial x} + \sum_{j=1}^{n} A_{ij}(x - y^2, y)\frac{\partial u_j}{\partial y} + \sum_{j=1}^{n} B_{ij}(x - y^2, y)\,u_j \tag{4.3}$$

or, after solving for the derivatives with respect to x and introducing new notations, we have

$$\frac{\partial u_i}{\partial x} = \sum_{j=1}^{n} a_{ij}(x, y)\frac{\partial u_j}{\partial y} + \sum_{j=1}^{n} b_{ij}(x, y)\, u_j \tag{4.4}$$
$$(i = 1, 2, \ldots, n).$$

(The possibility of solving the system (4.3) for these derivatives follows from the non-vanishing of the determinant of the system close to the origin O of the xy-plane. The reader should verify this.) The coefficients a_{ij} and b_{ij} are analytic close to the origin. The functions u_i are continuously differentiable close to the origin and they do not vanish on the parabola $y^2 = x$. We shall show that all the u_i are identically equal to zero close to the origin when $x > y^2$. This will automatically prove that all the u_i are identically equal to zero close to the origin when $x > 0$. The case in which $x < 0$ reduces to the case in which $x > 0$ by replacing x with $-x$.

Let us draw the straight line $x = a$ (for $a > 0$ (see Fig. 1)) and let us denote by H_a the region bounded by the segment l_a of this straight line and the portion K_a of the parabola $y^2 = x$. If a is sufficiently small, all the functions u_i and their first-order partial derivatives will be continuous up to the boundary of H_a. (We say that a function defined in a region H is continuous up to the boundary H^* of this region if it is possible to extend this function onto H^* in such a way that the extended function will be continuous in $H \cup H^*$.

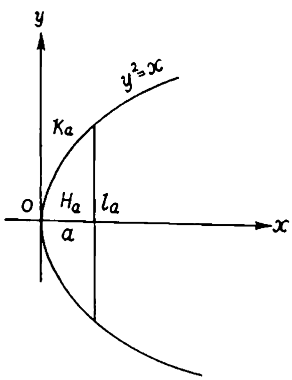


Fig. 1

For brevity, we introduce

$$F_i(u) \equiv \frac{\partial u_i}{\partial x} - \sum_{j=1}^{n} a_{ij} \frac{\partial u_j}{\partial y} - \sum_{j=1}^{n} b_{ij} u_j$$
$$(i = 1,\ 2,\ \ldots,\ n),$$

$$G_i(v)$$
$$\equiv \frac{\partial v_i}{\partial x} - \sum_{j=1}^{n} \frac{\partial}{\partial y}(a_{ji} v_j) + \sum_{j=1}^{n} b_{ji} v_j$$
$$(i = 1,\ 2,\ \ldots,\ n).$$

Suppose that two systems of functions u_i and v_i are defined in H_a and that these functions and their first partial derivatives are continuous up to the boundary of H_a. Then, integration by parts yields

$$\iint_{H_a} \left\{ \sum_{i=1}^{n} \left[v_i F_i(u) + u_i G_i(v) \right] \right\} dx\, dy$$
$$= \iint_{H_a} \left\{ \sum_{i=1}^{n} \left[\frac{\partial (u_i v_i)}{\partial x} - \sum_{j=1}^{n} \frac{\partial (a_{ij} u_j v_i)}{\partial y} \right] \right\} dx\, dy \tag{4.5}$$
$$= \int_{l_a} \sum_{i=1}^{n} u_i v_i dy + \int_{K_a} \sum_{i=1}^{n} u_i v_i dy + \int_{K_a} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} u_j v_i dx,$$

where we go around the curve bounding H_a (that is, $K_a \cup l_a$) in the positive direction. In particular, if u_i is the system of solutions of equations (4.4) that was defined above and if the system of $v_1, \ldots, v_n$ satisfies the equations

$$G_i(v) = 0 \qquad (i = 1,\ \ldots,\ n), \tag{4.6}$$

then we obtain from (4.5)

$$\int_{l_a} \sum_{i=1}^{n} u_i v_i dy = 0. \tag{4.7}$$

Now, we shall proceed to prove the Kovalevsky theorem for a system of linear equations. Let us use Remarks 1 and 2 to Section 2 and let us solve the system of equations (4.6) close to the point $(a,\ 0)$ giving as initial conditions on l_a all possible systems of polynomials. Since, for the constant

M, we may take the constant M^* defined in Remark 1 to Section 2 and general for all points $(a, 0)$, we may, on the basis of Remark 2 to Section 2, assert that if a is sufficiently small, all functions constituting a solution of the system (4.6) under such initial conditions will, as we know, be defined and analytic in H_a and, consequently, that they and their partial derivatives will be continuous up to the boundary of H_a.

Thus, equation (4.7) is valid if $0 < a < a_0$ (where a_0 is some fixed positive number) and all the v_i are arbitrary polynomials. Let us take any such a. Suppose that the length of the interval l_a is equal to s_a. From a well-known theorem of Weierstrass, for an arbitrary $\varepsilon > 0$, there is a system of polynomials v_i (for $i = 1, \ldots, n$) such that everywhere on l_a

$$|u_i - v_i| < \varepsilon \qquad (i = 1, \ldots, n). \tag{4.8}$$

By virtue of formulae (4.7) and (4.8),

$$\int_a \sum_{i=1}^{n} u_i^2 \, dy$$
$$= \int_{l_a} \sum_{i=1}^{n} u_i v_i dy + \int_{l_a} \sum_{i=1}^{n} u_i (u_i - v_i)\, dy \leqslant \varepsilon s_a \sum_{i=1}^{n} \max_{l_a} |u_i|,$$

so that, because of the arbitrariness of ε, we have

$$\int_{l_a} \sum_{i=1}^{n} u_i^2 \, dy = 0,$$

that is, all the u_i are identically equal to zero on l_a provided $0 < a < a_0$. This completes the proof of Holmgren's theorem.

With the help of this theorem and a change of independent variables, it is easy to prove the theorem of the uniqueness of the solution of the generalised Cauchy problem under the preceding hypotheses regarding the system (4.1) when the initial conditions are given on an analytic line that nowhere has a characteristic direction. The analogous theorem is valid also for linear systems with a greater number of independent variables when the initial conditions are given on an analytic surface S. Here, it is only necessary that this surface S nowhere has a characteristic direction. We may

require of the functions constituting a solution that they be given only on one side of S and that they and their first-order partial derivatives be continuous up to S. If under these conditions two solutions coincide on S, they will also coincide in some neighbourhood of S.

Remark: The region on the $\tilde{x}y$-plane in which the solution of the Cauchy problem for the system (4.1) is uniquely determined by the initial conditions (the region of uniqueness) can be described more precisely. Suppose that these initial conditions are given on the interval AB of the $\tilde{x}=0$-axis and that the solutions are being examined to the right of that axis. Through the points A and B we draw characteristics nearest to the interval AB but to the right of it. Then, we may show that the region bounded by the segment AB and these two characteristics will be a region of uniqueness of the solution of the Cauchy problem. The region of uniqueness when there is a greater number of independent variables is determined analogously (cf. Sections 10 and 12, in which the Cauchy problem is solved for hyperbolic equations).

2. Soon after the proof of Holmgren's theorem, Hadamard showed that the question of the uniqueness of the solution to the Cauchy problem for nonlinear equations close to S is easily reduced to the question of the uniqueness of the solution of the Cauchy problem for linear equations with sufficiently smooth though not necessarily analytic coefficients. Therefore, all subsequent efforts were concentrated on solving this last question. In 1938, Carleman solved it for systems of equations with partial derivatives with respect to two independent variables. Carleman's theorem is as follows:

Suppose that we are given the system of equations

$$\frac{\partial z_i}{\partial x}+\sum_{j=1}^{n} A_{ij}(x,\ y)\frac{\partial z_j}{\partial y}+\sum_{j=1}^{n} B_{ij}(x,\ y)\,z_j=0 \qquad (i=1,\ 2,\ \ldots,\ n) \tag{4.9}$$

that the functions A_{ij} and B_{ij} are defined in some closed region $\bar{G}$ of the half-plane $x\geqslant 0$ adjacent to the interval $|y|\leqslant a$ of the y-axis, that the A_{ij} have bounded first and second derivatives in $\bar{G}$, and that the B_{ij} are bounded in $\bar{G}$. Then, the solution of the system (4.9) in

$\bar{G}$ satisfying the conditions

$$z_i(0, y) = 0 \text{ when } |y| \leqslant a \quad (i = 1, \ldots, n)$$

and possessing continuous first derivatives with respect to x and y is identically equal to zero in some subregion $\bar{G}'$ of the region $\bar{G}$ that is adjacent to the interval $|y| \leqslant a$. Here, it is assumed that at every point in $\bar{G}$ all roots of the determinant

$$|A_{ij} - \lambda\delta_{ij}|$$

are distinct,* that is, that at no point of the region $\bar{G}$ are there coincident characteristic directions.

An analogous result for systems with many independent variables was obtained recently by Calderon**.

When the characteristic directions coincide, the uniqueness of the solution of the Cauchy problem may be destroyed. This was first shown by Myshkis (1947). As an example, he exhibited the system

$$\begin{aligned} \frac{\partial u}{\partial x} &= a_1(x, y)\frac{\partial u}{\partial y} + b_1(x, y)\frac{\partial v}{\partial y}, \\ \frac{\partial v}{\partial x} &= a_2(x, y)\frac{\partial u}{\partial y} + b_2(x, y)\frac{\partial v}{\partial y}, \end{aligned} \tag{4.10}$$

which has a solution u_0, v_0 such that the functions u_0 and v_0 possess continuous partial derivatives of all orders and are equal to zero on the straight line $x = 0$, but are nonzero arbitrarily close to the coordinate origin. In this case, the coefficients of the system are defined and differentiable on the entire plane, their derivatives have a discontinuity at $x = 0$, and the roots of the characteristic equation coincide on that line†.

In 1954, Pliś constructed a new example of a system of the form (4.10) that possesses a nontrivial solution of the Cauchy problem with zero initial conditions at $x = 0$ such

* $\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$

CALDERON, A.P., *American Journal of Mathematics*, **80, No. 1, 16-36 (1958).

†MYSHKIS, A.D., *Uspekhi matematicheskikh nauk*, **3**, No. 2, 3-46 (1948).

that the coefficients in the system have continuous partial derivatives of arbitrary order on the entire plane*.

The uniqueness of the solution of the Cauchy problem in the class of sufficiently smooth functions was proven for hyperbolic equations and hyperbolic systems with an arbitrary number of independent variables (we shall speak later of such systems) and also for a broad class of elliptic equations and systems (see Section 5). A vast literature has been devoted to the latter question.

This question of the uniqueness of the solution to the Cauchy problem in the class of non-analytic but sufficiently smooth functions is related to the question as to whether a sufficiently smooth real solution $(u_1, \ldots, u_N)$ of the system (3.13) of the preceding section defined in some real region of the space $(x_0, \ldots, x_n)$ on one side of and on a sufficiently smooth surface S that nowhere has a characteristic direction can be extended in a unique manner. Indeed, defining the functions u_i on one side of the surface S and on this surface itself determines the values on this surface of these functions u_i and their derivatives that appear in the Cauchy conditions. Thus, the question of extending the functions u_i to the other side of the surface S amounts to finding the solution of the generalised Cauchy problem in the region lying on the other side of the surface S. As was stated above, the question of the uniqueness of this solution has not as yet been completely clarified.

Similarly, the following question has also not yet been completely solved. Is it possible to extend in different ways a sufficiently smooth real solution $(u_1, \ldots, u_N)$ of the system (3.13) given in some real region of the space $(x_0, \ldots, x_n)$ lying on one side of a sufficiently smooth surface S and on that surface itself in the case in which the surface S is characteristic for the given system and the given solution? For all equations that we shall encounter, such an extension will always be possible by quite a large number of methods.

The question of the non-uniqueness of the extension of the solution of the system (3.13) beyond the characteristic is equivalent to the question of the existence of many solutions of the generalised Cauchy problem if the Cauchy conditions given on the characteristic are such that they generally admit at least one such solution. We have seen that for this case the functions u_i given on the characteristic and their

* PLIŚ, A., *Bull. Acad. Polon. Sci.*, 2, 55-57 (1954).

derivatives must in general satisfy certain relationships. These conditions are known to be satisfied if there exist functions $u_1, \ldots, u_N$ satisfying the given equations on one side or other of the characteristic.

If we are interested only in analytic solutions, the question of the uniqueness of the extension beyond the characteristic, or in general beyond any surface, of a given solution in an $(n+1)$-dimensional region can always be solved in the sense that such an extension is unique since an analytic function of $n+1$ independent variables is completely defined by its values in any arbitrarily small region of $(n+1)$-dimensional space.

3. We saw in subsection 3 of Section 3 that, if the surface S on which the Cauchy conditions are given has no characteristic direction anywhere, these Cauchy conditions together with the equations of the system (3.7) define uniquely the values on S of all the functions u_i and all their derivatives of order up to n_i. On the other hand, if the surface S has a characteristic close to the point A, the Cauchy conditions given on it admit different systems of values $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$, that may satisfy the system (3.7) if they admit at least one such system of values of $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$. (Here, we assume that the equation $\xi_0 = 0$ is an equation of the surface S.) Therefore, functions may exist that satisfy equations (3.7) everywhere in some region containing a portion of the characteristic surface, and the derivatives $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$ on this surface have a discontinuity of the first kind*. As these derivatives approach S from opposite sides, they approach different values, which satisfy simultaneously equations (3.7) on the surface S. If the surface S were that characteristic, the derivatives $\frac{\partial^{n_i} u_i}{\partial \xi_0^{n_i}}$ could not have discontinuities of the first kind on it when coefficients in equations (3.7) and all the other derivatives of the

* Translator's note. This is a common expression in Russian mathematical literature. A function is said to have a discontinuity of the first kind at a point if the left- and right-hand limits both exist as the argument approaches the point, but are not both equal to the value (if any) of the function at that point.

functions u_l of the form

$$\frac{\partial^k u_l}{\partial \xi_0^{k_0} \partial \xi_1^{k_1} \ldots \partial \xi_n^{k_n}},$$

$$k_0 + k_1 + \ldots + k_n = k \leqslant n_i, \quad k_0 < n_l$$

are continuous. Analogous statements are valid for nonlinear systems.

Example. Consider the equation

$$\frac{\partial^2 u}{\partial x \, \partial y} = 0, \tag{4.11}$$

for which the curves

$$x = \text{const}, \quad y = \text{const}$$

are characteristic.

Obviously equation (4.11) is satisfied by every function of the form

$$u = f(y),$$

where $f(y)$ is an arbitrary function that possesses a derivative everywhere. In particular, we may assume that the function $u = f(y)$ is such that its second derivative is everywhere continuous except at a single point, at which it possesses a discontinuity of the first kind. Then, we obtain a solution to equation (4.11) for which the second partial derivatives possess a discontinuity of the first kind on the characteristic.

In what remains we shall devote our attention primarily to equations of two types: Firstly, single second-order equations with a single unknown function, and, secondly, systems of equations of arbitrary order with an arbitrary number of unknown functions, but containing partial derivatives with respect to only two independent variables. These types of equations possess the property of being capable of reduction to certain simple 'canonical' forms. The methods of achieving this reduction will be described in the three sections which follow. The subject is divided into reduction to canonical form of single second-order equations with a single unknown function, single second-order partial differential equations with two independent variables, and systems of first-order partial differential equations with two independent variables.

5. REDUCTION TO CANONICAL FORM AT A POINT. CLASSIFICATION OF SECOND-ORDER EQUATIONS WITH ONE UNKNOWN FUNCTION

1. Consider the linear second-order equation

$$\sum_{i,j=1}^{n} A_{ij}(x_1, \ldots, x_n)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} B_i(x_1, \ldots, x_n)\frac{\partial u}{\partial x_i} + C(x_1, \ldots, x_n)u + F(x_1, \ldots, x_n) = 0 \quad (5.1)$$

With the single unknown function u, we assume $A_{ij} = A_{ji}$. All the functions A_{ij}, B_i, C, and F are real and are defined in some region G of the space $(x_1, \ldots, x_n)$.

We make a change of independent variables by setting

$$\xi_k = \sum_{i=1}^{n} a_{ki}x_i \qquad (k = 1, 2, \ldots, n), \quad (5.2)$$

where the a_{ki} are constants. We assume that the transformation (5.2) is nonsingular, that is, that the determinant $|a_{ki}|$ is nonzero. Then, the transformation from x_k to ξ_k is single-valued in both directions. Equation (5.1) can be written in terms of the independent variables $\xi_1, \xi_2, \ldots, \xi_n$ as follows*:

$$\sum_{k,l=1}^{n}\left(\sum_{i,j=1}^{n} A_{ij}a_{ki}a_{lj}\right)\frac{\partial^2 u}{\partial \xi_k \partial \xi_l} + \ldots = 0 \quad (5.3)$$

Here, we have written only the terms with the second-order derivatives of the unknown function u. It is clear from equation (5.3) that the coefficients of the second-order derivatives of u are changed when we make the change of independent variables indicated by formula (5.2) in just the same way as are the coefficients of the quadratic form

$$\sum_{i,j=1}^{n} A_{ij}x_i x_j \quad (5.4)$$

* To be sure of the validity of the change from the derivatives with respect to the independent variables x_i (for $i = 1, \ldots, n$) to derivatives with respect to the independent variables ξ_i (for $i = 1, \ldots, n$) according to the usual rules, it is sufficient to assume that the function u has continuous first and second derivatives.

in the change from x_k to ξ_k given by the formula

$$x_k = \sum_{i=1}^{n} a_{ik}\xi_i \qquad (k = 1, \ldots, n). \tag{5.5}$$

The coefficients A_{ij} of formula (5.4) are assumed to be constants equal to the values at some point $(x_1^0, \ldots, x_n^0)$ of the region G of the coefficients $A_{ij}(x_1, \ldots, x_n)$ in equation (5.1).

The existence of a real nonsingular transformation (5.5) that reduces every form (5.4) with real coefficients A_{ij} to the form

$$\sum_{i=1}^{m} \pm \xi_i^2, \text{where } m \leqslant n. \tag{5.6}$$

is shown in algebra. There exist many nonsingular real transformations (5.5) that reduce the form (5.4) to the form (5.6), but the number of terms with positive and the number with negative signs in the form (5.6) is determined exclusively by the form (5.4) and does not depend on the choice of nonsingular transformation (5.5). (This is the law of inertia of quadratic forms†.) The determinant $|A_{ik} - \lambda\delta_{ik}|$ will have only real roots λ. The number of terms in (5.6) with positive signs and the number of terms with negative signs will be equal to the number of positive and negative roots λ of this determinant respectively.

If we find some transformation (5.5) that reduces the form (5.4) to the form (5.6), the transformation (5.2) with the matrix that is the transpose and inverse of (a_{ik}) will reduce equation (5.1) to the form

$$\sum_{i,j=1}^{n} A_{ij}^{*}(x_1, \ldots, x_n) \frac{\partial^2 u}{\partial\xi_i \partial\xi_j} + \ldots = 0, \tag{5.7}$$

where

$$A_{ij}^{*}(x_1^0, \ldots, x_n^0) = \pm 1, \quad \text{if} \quad i = j \leqslant m,$$
$$A_{ij}^{*}(x_1^0, \ldots, x_n^0) = 0, \quad \text{if} \quad i \neq j \quad \text{or} \quad \text{if} \quad i = j > m.$$

Here, we have written only the terms with the derivatives of

† See KUROSH, A.G., *Course in higher algebra*, Fizmatgiz, Section 27 (1959), or GEL'FAND, I.M., Lectures on linear algebra, Gostekhizdat, 143 (1951).

the function u of highest order. The form (5.7) of equation (5.1) is called its canonical form at the point $(x_1^0, \dots, x_n^0)$.

Thus, for every point $(x_1^0, \dots, x_n^0)$ of the region G, it is possible to exhibit a nonsingular transformation (5.2) of the independent variables that reduces equation (5.1) to canonical form at that point.

For every point $(x_1^0, \dots, x_n^0)$, there is, generally speaking, a transformation (5.2) that reduces equation (5.1) to canonical form at that point. At other points, this transformation may not reduce the equation to canonical form. Examples show that when the number of independent variables is greater than 2, we not only cannot exhibit a linear transformation of the independent variables with constant coefficients but we cannot exhibit any other kind of nonsingular transformation of the variables that reduces the given linear second-order equation to canonical form even in an arbitrarily small region. On the other hand, in the case of two independent variables, such a transformation exists under quite broad assumptions regarding the coefficients in equation (5.1), as will be shown in the following subsection.

The classification of second-order equations is based on the possibility of reducing equation (5.1) to canonical form at a point.

2. Equation (5.1) is said to be elliptic at the point $(x_1^0, \dots, x_n^0)$ if all the $\overset{*}{A}_{ii}(x_1^0, \dots, x_n^0)$ (for $i = 1, \dots, n$) in equation (5.7) are nonzero and have the same sign.

Equation (5.1) is said to be hyperbolic at the point $(x_1^0, \dots, x_n^0)$ if all the $\overset{*}{A}_{ii}(x_1^0, \dots, x_n^0)$ have the same sign with the exception of one $\overset{*}{A}_{ii}$ that has the opposite sign and if $m = n$.

Equation (5.1) is said to be ultrahyperbolic at the point $(x_1^0, \dots, x_n^0)$ if equation (5.7) has more than one positive $\overset{*}{A}_{ii}(x_1^0, \dots, x_n^0)$ and more than one negative $\overset{*}{A}_{ii}(x_1^0, \dots, x_n^0)$ and if $m = n$.

Equation (5.1) is said to be parabolic in the broad sense at the point $(x_1^0, \dots, x_n^0)$ if some of the $\overset{*}{A}_{ii}(x_1^0, \dots, x_n^0)$ are identically equal to zero, that is, if $m < n$.

Equation (5.1) is said to be parabolic in the narrow sense or simply parabolic at the point $(x_1^0, \dots, x_n^0)$ if only one of the coefficients $\overset{*}{A}_{ii}(x_1^0, \dots, x_n^0)$ (let us assume that it is $\overset{*}{A}_{11}$) is equal to zero, if all the other $\overset{*}{A}_{ii}(x_1^0, \dots, x_n^0)$ have the same

sign, and if the coefficient of $\frac{\partial u}{\partial \xi_1}$ is nonzero.

Equation (5.1) is said to be elliptic, hyperbolic, ultrahyperbolic, etc. throughout the entire region G if it is respectively elliptic, hyperbolic, ultrahyperbolic, etc. at every point in the region G.

In applications, we sometimes encounter equations that are elliptic in one subregion G_1 of the region G in question but hyperbolic in another subregion G_2 of the region G. Such equations are said to be equations of mixed type. An example of such an equation is the Tricomi equation

$$y\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$$

in any region G containing points of the x-axis†.

3. Consider the nonlinear second-order equation

$$\Phi\left(x_1, \ldots, x_n, u, \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}, \ldots, \frac{\partial^2 u}{\partial x_i \partial x_j}, \ldots\right)=0$$

containing a single unknown function u. For a given solution $u^*(x_1, \ldots, x_n)$,this equation is said to be eliptic, hyperbolic, or parabolic in the broad sense at the point $(x_1^0, \ldots, x_n^0)$ or throughout a region G if the equation

$$\sum_{i,j=1}^{n} A_{ij}(x_1, \ldots, x_n)\frac{\partial^2 u}{\partial x_i \partial x_j}=0,$$

where

$$A_{ij}(x_1, \ldots, x_n)=\frac{\partial \Phi}{\partial\left\{\frac{\partial^2 u}{\partial x_i \partial x_j}\right\}}. \tag{5.8}$$

† Equations of mixed type were first investigated by Tricomi (see his article 'Sulle equazioni linrai; alle derivate parziali di secondo ordine, di tipo misto', *Rend. reale accad. lincei*, ser. 5, **14**, 134-247 (1923). Interest in such equations grew especially after their connection with problems of gas dynamics was discovered (see FRANKL', F.I., *Izv. Akad. Nauk. SSSR, seriya matem.*, 9, 121-143 (1945)). A great deal of study has been made of equations of mixed type in recent years (see BITSADZE, A.V., *Equations of mixed type*, Izdat. Akad. Nauk. SSSR (1959), which includes a detailed bibliography of the question).

is elliptic, hyperbolic, or parabolic in the broad sense at the point $(x_1^0, \ldots, x_n^0)$, or throughout the region G respectively. On the right side of equation (5.8), the function $u^*(x_1, \ldots, x_n)$ and its derivatives have been substituted for the function u and its corresponding derivatives.

In what follows, we shall study only linear second-order equations with one unknown function. We shall also confine ourselves to equations that are either elliptic, hyperbolic, or parabolic throughout the entire region in question. We shall not deal with ultrahyperbolic equations. Such equations are not encountered either in physics or in technology. Also, we shall not deal with equations that are parabolic in the broad but not in the narrow sense. In accordance with this, when we discuss parabolic equations in Chapter IV, we shall always be referring only to equations that are parabolic in the narrow sense.

6. REDUCTION TO CANONICAL FORM OF SECOND-ORDER PARTIAL DIFFERENTIAL EQUATIONS IN TWO INDEPENDENT VARIABLES NEAR A POINT

1. Consider the equation†

$$A\frac{\partial^2 u}{\partial x^2} + 2B\frac{\partial^2 u}{\partial x\,\partial y} + C\frac{\partial^2 u}{\partial y^2} + F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0 \tag{6.1}$$

where the coefficients A, B, and C are functions of x and y that have continuous first and second derivatives. We shall assume that A, B, and C do not vanish simultaneously and that the function $u(x, y)$ has continuous first and second derivatives. Let us change from the independent variables x and y to the independent variables ξ and η. Suppose that

$$\xi = \xi(x, y), \quad \eta = \eta(x, y) \tag{6.2}$$

are twice continuously differentiable functions and that the Jacobian

$$\begin{vmatrix} \partial\xi/\partial x & \partial\xi/\partial y \\ \partial\eta/\partial x & \partial\eta/\partial y \end{vmatrix}$$

† In this section, we shall examine equations of a somewhat more general type than linear equations since all the considerations involved in reducing a linear equation to canonical form are likewise applicable for these equations.

does not vanish anywhere in the region in question. Then, the system (6.2) can be solved uniquely for x and y in some region of the ξ, η-plane. The functions $x(\xi, \eta)$ and $y(\xi, \eta)$ that we obtain will also be twice continuously differentiable functions of ξ and η. Equation (6.1) is then written in the new variables ξ and η as follows:

$$\begin{aligned}\frac{\partial^2 u}{\partial \xi^2}&\left[A\left(\frac{\partial \xi}{\partial x}\right)^2+2B\frac{\partial \xi}{\partial x}\frac{\partial \xi}{\partial y}+C\left(\frac{\partial \xi}{\partial y}\right)^2\right]\\ &+2\frac{\partial^2 u}{\partial \xi\,\partial \eta}\left[A\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x}+B\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial y}+B\frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial x}+C\frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y}\right]\\ &+\frac{\partial^2 u}{\partial \eta^2}\left[A\left(\frac{\partial \eta}{\partial x}\right)^2+2B\frac{\partial \eta}{\partial x}\frac{\partial \eta}{\partial y}+C\left(\frac{\partial \eta}{\partial y}\right)^2\right]\\ &\qquad\qquad+F_1\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial \eta}\right)=0.\end{aligned}\tag{6.3}$$

Let us show that, in some neighbourhood G of a fixed point (x_0, y_0), functions $\xi(x, y)$ and $\eta(x, y)$ can be chosen such that equation (6.3) will be in canonical form at every point of this neighbourhood. We need to examine separately the case in which $B^2 > AC$ (or $B^2 < AC$) at the point in question and the case in which $B^2 \equiv AC$ throughout some neighbourhood of that point. We shall not consider cases in which the expression $B^2 - AC$ changes sign or vanishes (other than identically) in an arbitrary neighbourhood of the point in question.

2. Let us consider first the case in which $B^2 > AC$ throughout the entire region in question, that is, the case in which equation (6.1) is hyperbolic (see the definition in the preceding section). We may assume that either $A \neq 0$ or $C \neq 0$ at the point (x_0, y_0) in a neighbourhood of which we shall reduce equation (6.1) to canonical form. In the opposite case, we would be able to obtain such a situation by a change of variables as follows:

$$x = x' + y',$$
$$y = x' - y'.$$

Consider the equation

$$A\left(\frac{\partial \varphi}{\partial x}\right)^2+2B\frac{\partial \varphi}{\partial x}\frac{\partial \varphi}{\partial y}+C\left(\frac{\partial \varphi}{\partial y}\right)^2=0. \tag{6.4}$$

Suppose $A \neq 0$. As $B^2 - AC > 0$, equation (6.4) can be written

in the form

$$\left[A\frac{\partial\varphi}{\partial x}-(-B-\sqrt{B^2-AC})\frac{\partial\varphi}{\partial y}\right]\times\left[A\frac{\partial\varphi}{\partial x}-(-B+\sqrt{B^2-AC})\frac{\partial\varphi}{\partial y}\right]=0,$$

and therefore equation (6.4) is satisfied by the solutions of each of the equations

$$\left.\begin{aligned}A\frac{\partial\varphi_1}{\partial x}&=(-B-\sqrt{B^2-AC})\frac{\partial\varphi_1}{\partial y},\\ A\frac{\partial\varphi_2}{\partial x}&=(-B+\sqrt{B^2-AC})\frac{\partial\varphi_2}{\partial y}.\end{aligned}\right\}\tag{6.5}$$

Let us determine the functions $\varphi_i(x, y)$ (for $i=1, 2$) as solutions to equations (6.5) by giving their values respectively on certain curves l_i (for $i=1, 2$) passing through the point (x_0, y_0) that are nowhere tangent to the characteristics of the corresponding equation. If the curves l_i and the values of the functions φ_i that are given on them are chosen sufficiently smooth, we shall obtain solutions $\varphi_i(x, y)$ (for $i=1, 2$) that have continuous first and second derivatives with respect to x and y. If we assume also that the initial values $\varphi_i(x, y)$ on the l_i are chosen in such a way that the derivative φ_i in the direction l_i does not vanish at the point (x_0, y_0), then the partial derivatives of the function $\varphi_i(x, y)$ with respect to x and y cannot both vanish at that point (otherwise, the derivative at that point in an arbitrary direction would vanish)*.

Since $A\neq 0$, it follows from equations (6.5) that in this case $\frac{\partial\varphi_1}{\partial y}\neq 0$ and $\frac{\partial\varphi_2}{\partial y}\neq 0$ in a neighbourhood of the point (x_0, y_0) and that

$$\frac{\partial\varphi_1}{\partial x}:\frac{\partial\varphi_1}{\partial y}=\frac{-B-\sqrt{B^2-AC}}{A};$$

$$\frac{\partial\varphi_2}{\partial x}:\frac{\partial\varphi_2}{\partial y}=\frac{-B+\sqrt{B^2-AC}}{A}.$$

* See Section 53 of my *Lectures on the theory of ordinary differential equations* (1952). I draw attention to the fact that, in the case of two independent variables, the definition of characteristics that we employed in Section 3 of the present book coincides with the definition of characteristics given in Section 53 of the cited work. In the case of a larger number of independent variables, however, these definitions are completely different.

Since $B^2 - AC \neq 0$, we have

$$\frac{\partial \varphi_1}{\partial x} : \frac{\partial \varphi_1}{\partial y} \neq \frac{\partial \varphi_2}{\partial x} : \frac{\partial \varphi_2}{\partial y}.$$

It follows from this that the Jacobian

$$J = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x} & \frac{\partial \varphi_1}{\partial y} \\ \frac{\partial \varphi_2}{\partial x} & \frac{\partial \varphi_2}{\partial y} \end{vmatrix} \tag{6.6}$$

is nonzero in some neighbourhood G of the point (x_0, y_0). Therefore, in this neighbourhood we may take

$$\left.\begin{aligned} \xi &= \xi(x, y) = \varphi_1(x, y), \\ \eta &= \eta(x, y) = \varphi_2(x, y). \end{aligned}\right\} \tag{6.7}$$

in equations (6.2).

Then, the terms containing $\frac{\partial^2 u}{\partial \xi^2}$ and $\frac{\partial^2 u}{\partial \eta^2}$ vanish. The coefficients of $\frac{\partial^2 u}{\partial \xi\, \partial \eta}$ will then be nonzero throughout the entire region G in question, since otherwise the order of the equation would be decreased in the change from the coordinates (x, y) to the coordinates (ξ, η), and, consequently, in the reverse change from the coordinates (ξ, η) to the coordinates (x, y), the order of the equation would be increased at some point, which obviously is impossible.

If we divide equation (6.3) by the coefficient of $\frac{\partial^2 u}{\partial \xi\, \partial \eta}$, we reduce this equation to the form

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta} = F_2\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial \eta}\right) \tag{6.8}$$

in a neighbourhood G of the point (x_0, y_0). This form of the equation is also called canonical.

If we set $\xi = \alpha + \beta$ and $\eta = \alpha - \beta$, we reduce equation (6.8) to the form

$$\frac{\partial^2 u}{\partial \alpha^2} - \frac{\partial^2 u}{\partial \beta^2} = \Phi\left(\alpha, \beta, u, \frac{\partial u}{\partial \alpha}, \frac{\partial u}{\partial \beta}\right). \tag{6.9}$$

After reducing a hyperbolic equation to the canonical form (6.8), it is sometimes possible to integrate it in closed form, that is, to find a formula giving all solutions of the equation.

Example 1. The equation

$$\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}=0 \tag{6.10}$$

can, by the change of independent variables

$$x=\frac{\xi+\eta}{2}, \quad y=\frac{\xi-\eta}{2}$$

be reduced to the form

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta}=0. \tag{6.11}$$

If we denote $\frac{\partial u}{\partial \eta}$ by v, we obtain $\frac{\partial v}{\partial \xi}=0$, so that $v=f(\eta)$, where f is an arbitrary function of η. If we consider ξ in the equation

$$\frac{\partial u}{\partial \eta}=f(\eta)$$

as a parameter and integrate this equation, we obtain

$$u=\int f(\eta)\,d\eta+C(\xi)$$

or

$$u=\varphi(\xi)+\psi(\eta)=\varphi(x+y)+\psi(x-y), \tag{6.12}$$

where φ and ψ are twice continuously differentiable functions.

Example 2. The equation

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta}=\frac{1}{2\xi}\frac{\partial u}{\partial \eta} \qquad (\xi \neq 0) \tag{6.13}$$

is, after we make the change

$$\frac{\partial u}{\partial \eta}=v$$

reduced to the equation

$$\frac{\partial v}{\partial \xi}=\frac{1}{2\xi}v.$$

This equation is easily integrated by the method of separation of variables. Since η appears in v as a parameter,

the constant of integration will be a function of that parameter. We obtain

$$\ln|v| = \frac{1}{2}\ln|\xi| + \ln|C(\eta)|$$

or

$$v = \frac{\partial u}{\partial \eta} = C(\eta)\sqrt{|\xi|}.$$

Therefore,

$$u = C_1(\eta)\sqrt{|\xi|} + C_2(\xi).$$

Here,

$$C_1(\eta) = \int C(\eta)\, d\eta$$

is an arbitrary differentiable function of η (since $C(\eta)$ is arbitrary) and $C_2(\xi)$ is an arbitrary function of ξ.

3. If

$$B^2 = AC$$

throughout the entire region in question, equation (6.1) will be a parabolic equation in this region (see the definition of a parabolic equation in the preceding section). We assume that the coefficients A, B, and C in equation (6.1) do not simultaneously vanish in the region in question. Since $B^2 = AC$, it follows that at every point in the region one of the coefficients A and C is nonzero. Suppose, for example, that $A \neq 0$ at a particular point (x_0, y_0). Then, the two equations (6.5) coincide and become the equation

$$A\frac{\partial \varphi}{\partial x} + B\frac{\partial \varphi}{\partial y} = 0. \qquad (6.14)$$

Every solution to equation (6.14) must, by virtue of the condition $B^2 = AC$, also satisfy the equation

$$B\frac{\partial \varphi}{\partial x} + C\frac{\partial \varphi}{\partial y} = 0. \qquad (6.15)$$

Just as in the preceding subsection we may find a solution $\varphi(x, y)$ of equation (6.14) such that the function $\varphi(x, y)$ has

continuous second-order derivatives and such that its first derivatives do not simultaneously vanish in some neighbourhood G of the point (x_0, y_0). We may assume that $A \neq 0$ throughout the entire region G.

Suppose that

$$\psi(x, y) = \text{const}$$

is a family of curves in the region G such that the function $\psi(x, y)$ has continuous second-order derivatives and that the Jacobian

$$J = \begin{vmatrix} \frac{\partial\varphi}{\partial x} & \frac{\partial\varphi}{\partial y} \\ \frac{\partial\psi}{\partial x} & \frac{\partial\psi}{\partial y} \end{vmatrix} \tag{6.16}$$

does not vanish anywhere in the region G. (Since $A \neq 0$ and consequently $\frac{\partial\varphi}{\partial y} \neq 0$ in G, we may, for example, take $\psi(x, y) \equiv x$.) In formulae (6.2), we set

$$\xi = \varphi(x, y) \quad \text{and} \quad \eta = \psi(x, y).$$

Then, the coefficient of $\frac{\partial^2 u}{\partial \xi^2}$ in equation (6.3) vanishes and the coefficient of $\frac{\partial^2 u}{\partial \xi\, \partial \eta}$ becomes

$$\left(A\frac{\partial\varphi}{\partial x} + B\frac{\partial\varphi}{\partial y}\right)\frac{\partial\psi}{\partial x} + \left(B\frac{\partial\varphi}{\partial x} + C\frac{\partial\varphi}{\partial y}\right)\frac{\partial\psi}{\partial y}.$$

According to (6.14) and (6.15), it will also be identically equal to zero in the region G.

The coefficient of $\frac{\partial^2 u}{\partial \eta^2}$ in equation (6.3) takes the form

$$A\left(\frac{\partial\psi}{\partial x}\right)^2 + 2B\frac{\partial\psi}{\partial x}\frac{\partial\psi}{\partial y} + C\left(\frac{\partial\psi}{\partial y}\right)^2 = \frac{1}{A}\left(A\frac{\partial\psi}{\partial x} + B\frac{\partial\psi}{\partial y}\right)^2.$$

This expression cannot vanish since, if it did, the Jacobian (6.16) would, on the basis of (6.14), vanish in the region G. Therefore, equation (6.3) can be divided by this coefficient. We then have

$$\frac{\partial^2 u}{\partial \eta^2} = F_2\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial \eta}\right). \tag{6.17}$$

This equation is in canonical form in the region G; canonical form was defined in Section 5.

If equation (6.1) is linear, then equation (6.17) will also be linear. Suppose that it is of the form

$$\frac{\partial^2 u}{\partial \eta^2} = A_1 \frac{\partial u}{\partial \xi} + B_1 \frac{\partial u}{\partial \eta} + C_1 u + D_1. \tag{6.18}$$

We can still simplify this equation somewhat by replacing u with a new unknown function z. Let us set

$$u = zv,$$

where $v(\xi, \eta)$ is a function of ξ and η, which we shall determine below. Then, equation (6.18) is replaced with the equation

$$v \frac{\partial^2 z}{\partial \eta^2} + 2 \frac{\partial v}{\partial \eta} \frac{\partial z}{\partial \eta} = A_1 v \frac{\partial z}{\partial \xi} + B_1 v \frac{\partial z}{\partial \eta} + C_2 z + D_1. \tag{6.19}$$

Here, we have written separately only the terms containing derivatives of z. All the terms containing the function z itself are included in $C_2 z$. Let us choose a function $v(\xi, \eta)$ such that the derivative $\frac{\partial z}{\partial \eta}$ vanishes in equation (6.19). If we set the coefficient of $\frac{\partial z}{\partial \eta}$ equal to zero, we obtain

$$\frac{\partial^2 z}{\partial \eta^2} = A_1 \frac{\partial z}{\partial \xi} + C_3 z + D_2, \tag{6.20}$$

where

$$C_3 = \frac{C_2}{v}, \quad D_2 = \frac{D_1}{v}, \quad v(\xi, \eta) = e^{\frac{1}{2}\int B_1(\xi, \eta)\, d\eta}.$$

4. Finally, let us examine the case in which

$$AC > B^2$$

everywhere in the region in question. Then, equation (6.1) will be an elliptic equation in that region (see the definition of an elliptic equation in Section 5). In this case, we shall assume that all the coefficients A, B and C are analytic functions of x and y. Then, the coefficients in equations (6.5) are also analytic functions of x and y. Suppose that

$$\varphi(x, y) = \varphi^*(x, y) + i\varphi^{**}(x, y)$$

is an analytic solution of the first of equations (6.5) in a neighbourhood of the point† (x_0, y_0) and suppose that $\left|\frac{\partial\varphi}{\partial x}\right| + \left|\frac{\partial\varphi}{\partial y}\right| \neq 0$ in this neighbourhood.

In equations (6.2), let us set

$$\xi = \varphi^* (x, y) \text{ and } \eta = \varphi^{**} (x, y). \tag{6.21}$$

Equations (6.21) can be solved for x and y since the Jacobian

$$J = \begin{vmatrix} \frac{\partial\xi}{\partial x} & \frac{\partial\xi}{\partial y} \\ \frac{\partial\eta}{\partial x} & \frac{\partial\eta}{\partial y} \end{vmatrix} \tag{6.22}$$

does not vanish anywhere. To show this, we separate the real and imaginary parts in equations (6.5) and obtain

$$\left.\begin{aligned} A\frac{\partial\xi}{\partial x} &= -B\frac{\partial\xi}{\partial y} + \sqrt{AC-B^2}\,\frac{\partial\eta}{\partial y}, \\ A\frac{\partial\eta}{\partial x} &= -B\frac{\partial\eta}{\partial y} - \sqrt{AC-B^2}\,\frac{\partial\xi}{\partial y}. \end{aligned}\right\} \tag{6.23}$$

Substituting the expressions obtained for $\frac{\partial\xi}{\partial x}$ and $\frac{\partial\eta}{\partial x}$ in the Jacobian (6.22), we obtain

$$J = \frac{\sqrt{AC-B^2}}{A}\left[\left(\frac{\partial\xi}{\partial y}\right)^2 + \left(\frac{\partial\eta}{\partial y}\right)^2\right].$$

From this, it is clear that this determinant can be equal to zero only at those points at which

$$\frac{\partial\xi}{\partial y} = \frac{\partial\eta}{\partial y} = 0,$$

and hence, on the basis of equations (6.23), only at points at which

† It is possible to find an analytic solution $\phi(x, y)$ of equation (6.5) such that $\partial\phi/\partial x$ and $\partial\phi/\partial y$ do not simultaneously vanish in some neighbourhood of an arbitrary point (x_0, y_0) of the region in question. For example, this may be done by giving values of $\phi(x, y)$ for $x = x_0$ such that $\phi'_y(x_0, y_0) \neq 0$ in accordance with the Kovalevsky theorem. We assume that $\phi(x, y)$ is such a solution.

$$\frac{\partial \xi}{\partial x}=0 \text{ and } \frac{\partial \eta}{\partial x}=0.$$

But such points do not exist in the region in question since we would then have at such points

$$\frac{\partial \varphi}{\partial x}=\frac{\partial \varphi}{\partial y}=0.$$

If we separate the real and imaginary parts in the identity

$$A\left(\frac{\partial \varphi}{\partial x}\right)^2+2B\frac{\partial \varphi}{\partial x}\frac{\partial \varphi}{\partial y}+C\left(\frac{\partial \varphi}{\partial y}\right)^2=0$$

we obtain

$$A\left(\frac{\partial \xi}{\partial x}\right)^2+2B\frac{\partial \xi}{\partial x}\frac{\partial \xi}{\partial y}+C\left(\frac{\partial \xi}{\partial y}\right)^2 = A\left(\frac{\partial \eta}{\partial x}\right)^2+2B\frac{\partial \eta}{\partial x}\frac{\partial \eta}{\partial y}+C\left(\frac{\partial \eta}{\partial y}\right) \tag{6.24}$$

$$A\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x}+B\left(\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial y}+\frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial x}\right)+C\frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y}=0. \tag{6.25}$$

Because of the positive definiteness of the form

$$A\alpha^2+2B\alpha\beta+C\beta^2 \qquad (B^2-AC<0),$$

the right and left sides of equation (6.24) can vanish only when

$$\frac{\partial \xi}{\partial x}=\frac{\partial \xi}{\partial y}=\frac{\partial \eta}{\partial x}=\frac{\partial \eta}{\partial y}=0. \tag{6.26}$$

But we chose the function $\varphi(x, y)$ in such a way that equations (6.26) are not simultaneously satisfied. Thus, the coefficients of $\frac{\partial^2 u}{\partial \xi^2}$ and $\frac{\partial^2 u}{\partial \eta^2}$ in equation (6.3) coincide and do not vanish. Therefore, equation (6.3) is reduced to the form

$$\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial \eta^2}=F_2\left(\xi, \eta, u, \frac{\partial u}{\partial \xi}, \frac{\partial u}{\partial \eta}\right). \tag{6.27}$$

This form of an elliptic equation is called its canonical form.

We have reduced the equation to such a form in a neighbourhood of some point (x_0, y_0) at which there exists an analytic solution of equations (6.5) with nonvanishing derivatives. By other and more complicated reasoning, one can

show that such reduction to canonical form is possible without assuming analyticity of $A(x, y)$, $B(x, y)$, and $C(x, y)$ but assuming only that these functions possess continuous first and second derivatives.

7. REDUCTION TO CANONICAL FORM OF A SYSTEM OF LINEAR EQUATIONS WITH SECOND-ORDER PARTIAL DERIVATIVES WITH RESPECT TO TWO INDEPENDENT VARIABLES

Consider the system of equations

$$\frac{\partial u_i}{\partial x}=\sum_{j=1}^{n} a_{ij}(x, y)\frac{\partial u_j}{\partial y}+f_i(x, y, u_1, \ldots, u_n) \qquad (7.1)$$
$$(i=1, 2, \ldots, n).$$

It is immaterial to us whether the f_i are linear or nonlinear with respect to the $u_1, u_2, \ldots, u_n$.We shall assume that the coefficients $a_{ij}(x, y)$ are real and that they have continuous partial derivatives with respect to x and y up to the k-th inclusively (where $k \geqslant 1$) in some region G in the xy-plane. Then, under certain addition hypotheses which we shall later write in italics, the system (7.1) can, in some neighbourhood of an arbitrarily chosen point A within G, be reduced to the canonical form

$$\begin{aligned}
\frac{\partial v_1}{\partial x}&=\lambda_1(x, y)\frac{\partial v_1}{\partial y} +f_1^*(x, y, v_1, \ldots, v_n), \qquad (7.2)\\
\frac{\partial v_2}{\partial x}&=\alpha_1(x, y)\frac{\partial v_1}{\partial y}+\lambda_1(x, y)\frac{\partial v_2}{\partial y} +f_2^*(x, y, v_1, \ldots, v_n),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\frac{\partial v_{n_1}}{\partial x}&=\alpha_{n_1-1}(x, y)\frac{\partial v_{n_1-1}}{\partial y}+\lambda_1\frac{\partial v_{n_1}}{\partial y}+f_{n_1}^*(x, y, v_1, \ldots, v_n),\\
\frac{\partial v_{n_1+1}}{\partial x}&=\lambda_2(x, y)\frac{\partial v_{n_1+1}}{\partial y}+f_{n_1+1}^*(x, y, v_1, \ldots, v_n),\\
\frac{\partial v_{n_1+2}}{\partial x}&=\beta_1(x, y)\frac{\partial v_{n_1+1}}{\partial y}+\lambda_2(x, y)\frac{\partial v_{n_1+2}}{\partial y}\\
&\qquad +f_{n_1+2}^*(x, y, v_1, \ldots, v_n),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\frac{\partial v_{n_1+n_2}}{\partial x}&=\beta_{n_2-1}(x, y)\frac{\partial v_{n_1+n_2-1}}{\partial y}+\lambda_2(x, y)\frac{\partial v_{n_1+n_2}}{\partial y}\\
&\qquad +f_{n_1+n_2}^*(x, y, v_1, \ldots, v_n),
\end{aligned}$$

. . . .

$$\frac{\partial v_{n-n_k+1}}{\partial x}=\lambda_k(x, y)\frac{\partial v_{n-n_k+1}}{\partial y}+f^*_{n-n_k+1}(x, y, v_1,\ldots, v_n),$$

$$\frac{\partial v_{n-n_k+2}}{\partial x}=\omega_1(x, y)\frac{\partial v_{n-n_k+1}}{\partial y}+\lambda_k(x, y)\frac{\partial v_{n-n_k+2}}{\partial y}+ \\ +f^*_{n-n_k+2}(x, y, v_1,\ldots, v_n),$$

. . . .

$$\frac{\partial v_n}{\partial x}=\omega_{n_k-1}(x,y)\frac{\partial v_{n-1}}{\partial y}+\lambda_k(x, y)\frac{\partial v_n}{\partial y} \\ +f^*_n(x, y, v_1, \ldots, v_n).$$

by a linear nonsingular transformation of the unknown functions $u_1, \ldots, u_n$ with coefficients possessing as many continuous derivatives as do the coefficients $a_{ij}(x, y)$. Here, $\lambda_1(x, y), \ldots, \lambda_k(x, y)$ are the roots of the determinant of the matrix

$$\| a_{ij}(x, y) \| - \lambda E, \tag{7.3}$$

and $\alpha_i(x, y), \beta_i(x, y), \ldots, \omega_i(x, y)$ are certain rather arbitrary functions possessing continuous derivatives up to the k-th order inclusively and nowhere vanishing in the neighbourhood in question of the point A. The functions $v_i, \lambda_i, \alpha_i, \beta_i, \ldots, \omega_i, f^*_1, \ldots, f^*_n$ can in general be complex functions of their arguments. If $f_1, \ldots, f_n$ have continuous derivatives of q-th order, then $f^*_1, \ldots, f^*_n$ will have continuous derivatives up to the order $\min\{q, k-1\}$ inclusively.

The systems (7.1) and (7.2) that we have been considering differ from a system of linear ordinary differential equations

$$\frac{dy_i}{dx}=\sum_{j=1}^{n} a_{ij}y_j+f_i(x) \qquad (i=1, 2,\ldots, n) \tag{7.4}$$

with constant coefficients a_{ij} and the canonical system (1.33) corresponding to it shown in Section 43 of my book on ordinary differential equations (Gostekhizdat, 1952) only in the fact that instead of $\frac{\partial}{\partial x}$ on the left sides of the corresponding ordinary differential equations we have $\frac{d}{dx}$, and instead of $\frac{\partial}{\partial y}$, in the corresponding ordinary differential equations we

understand the factor 1. The coefficients are constant in the system of ordinary differential equations (133) and the functions f and f^* depend only on a single independent variable whereas, in the corresponding partial differential equations that we have been examining, the coefficients of the derivatives depend on two independent variables and the functions f and f^* depend not only on these two independent variables but also on all the unknown functions.

The reduction of the system (7.1) to the canonical form (7.2) is carried out by a change in the unknown functions in just the same way as was done in Section 44 in my book on ordinary differential equations for a system of linear equations with constant coefficients. The only thing that we need to worry about in this case is to show that close to the point A the coefficients of the linear transformation described in Section 44 are functions (of x and y) possessing the same degree of smoothness as do the coefficients $a_{ij}(x, y)$ of the system (7.1). For this, it may be advisable for us to review somewhat this Section 44. We shall use the method of mathematical induction. For $n=1$, the assertion that we made regarding the possibility of reducing the system (7.1) to the form (7.2) by a linear transformation with smooth coefficients is obvious. Let us assume that it is valid for a set of $n-1$ equations. Let us show that it is then valid also for a set of n equations.

We multiply the i-th equation of the system (7.1) by k_i, where the k_i are certain differentiable functions in a neighbourhood of the point A; we shall determine these functions later. Let us sum the equations thus obtained over all the i and let us write the result in the form

$$\frac{\partial\left(\sum\limits_i k_i u_i\right)}{\partial x}$$

$$=\sum_{i,j}\frac{\partial(a_{ij}k_iu_j)}{\partial y}+\sum_i k_if_i+\sum_i u_i\frac{\partial k_i}{\partial x}-\sum_{i,j}u_j\frac{\partial(a_{ij}k_i)}{\partial y}.$$

Now, let us determine the k_i in such a way that

$$\sum_{i,j}a_{ij}k_iu_j\equiv\lambda\sum_i k_iu_i, \tag{7.5}$$

for all u_j, where λ is a certain differentiable function of x and y, either real or complex. For this, it is sufficient and

clearly necessary that the coefficients of u_j on the two sides of this identity be the same, that is, that

$$\lambda k_j = \sum_i a_{ij} k_i, \quad j = 1, 2, \ldots, n. \tag{7.6}$$

Thus, to determine $k_1, k_2, \ldots, k_n$, we obtain a system of n linear homogeneous equations with n unknowns. For this system to have a nontrivial solution (which is the only kind that we are interested in), it is necessary and sufficient that the determinant of its coefficient be equal to zero. This condition can be written as follows:

$$|\lambda E - \| a_{ij} \| \,| = 0. \tag{7.7}$$

The matrix $\lambda E - \| a_{ij} \|$ is called the characteristic matrix of the system (7.1).

Suppose that λ_1 is one of the roots of equation (7.7). Let us assume that in the neighbourhood in question of the point A, every root of equation (7.7) has the same multiplicity for all points in that neighbourhood. Suppose that λ_1 is of multiplicity α_1 in that neighbourhood. Then, in that neighbourhood, λ_1 satisfies the algebraic equation

$$f^{(\alpha_1 - 1)}(\lambda, x, y) = 0,$$

where $f^{(k)}(\lambda, x, y)$ is the kth-order derivative of the left side of equation (7.7) with respect to λ. Then, throughout this neighbourhood, $f^{(\alpha_1)}(\lambda_1(x, y), x, y) \neq 0$. Therefore, on the basis of the well-known theorem on implicit functions, $\lambda_1(x, y)$ will be equally smooth in the neighbourhood of the point A; that is, it will have as many continuous derivatives with respect to x and y as do the coefficients a_{ij}.

We assume also, that in the neighbourhood in question of the point A, the matrix

$$\| a_{ij} \| - \lambda_k E, \tag{7.8}$$

where λ_k is a root of equation (7.7), has the same rank r_k*. Then, in this neighbourhood of the point A, the system (7.6)

*It is easy to show that $r_k > n - \alpha_k$. Specifically, the derivative of order α_k of the determinant (7.7) with respect to λ is, as is easily seen, a linear combination of the minors of order $(n - \alpha_k)$ of the determinant (7.8) when $\lambda = \lambda_k$. Since this derivative is nonzero, one of the minors order $(n - \alpha_k)$ of the matrix (7.8) is nonzero.

has a solution with $\lambda=\lambda_1$ consisting of functions that do not vanish simultaneously anywhere in the neighbourhood of the point A. Furthermore, these functions have the same smoothness as do the a_{ij}. Let us denote them by k_{1i}. To find out what the k_{1i} are, we note the following. Since the rank of the matrix (7.8) is everywhere r_1 in the neighbourhood of A, this point A has a neighbourhood in which some particular $n-r_1$ equations of the system (7.6) are consequences of the remaining r_1 equations. Therefore, every system of functions k_{1i} that satisfy these r_1 equations in some small neighbourhood of the point A will satisfy the entire system (7.6). To find the solution of these r_1 equations (for brevity, we shall call them the C_1 equations), we note the following. Since the rank of the matrix (7.8) for $\lambda=\lambda_1$ is equal to r_1, it is possible to construct a square matrix with nonzero determinant in some neighbourhood of the point A from the columns of the matrix composed of the coefficients of the C_1 system. The functions k_{1i}, which are factors in these columns, will be treated as unknowns. The remaining k_{1i} functions we set equal to arbitrary constants that are not all equal to zero. For definiteness, we set them all equal to 1. Then, the system C_1 determines uniquely all the remaining k_{1j} as functions having the same smoothness as do the a_{ij}.

Thus, we have found functions k_{1i} (for $i=1, 2, \ldots, n$) that do not vanish simultaneously anywhere in some neighbourhood of the point A and that have the same smoothness as do the a_{ij}. For definiteness, let us assume that $k_{11}\neq 0$ at the point A. Obviously, this in no way restricts the generality since we can always achieve this by a change in the numbering of the u_i, which reduces to a nonsingular linear transformation of the u_i. Furthermore, let us set

$$z_1=\sum_{j=1}^{n} k_{1j}u_j.$$

Obviously, the function $z_1(x, y)$ satisfies the equation

$$\frac{\partial z_1}{\partial x}=\lambda_1\frac{\partial z_1}{\partial y}+f_1^*(x, y, z_1, u_2, \ldots, u_n),$$

where

$$f_1^*(x, y, z_1, u_2, \ldots, u_n)$$

$$\equiv \sum_i k_{1i} f_i + \sum_i u_i \frac{\partial k_{1i}}{\partial x} - \sum_{i,j} u_j \frac{\partial (a_{ij} k_{1i})}{\partial y} + z_1 \frac{\partial \lambda_1}{\partial y}$$

(see formula (7.5) and the equation preceding it).

Furthermore, all the considerations made in Section 44 in my book on ordinary differential equations are applicable without any significant changes†. This reasoning is considerably simplified in the case in which all the roots λ of equation (7.7) are distinct, and we shall follow it to the end. In this case, to every root λ_i (for $i = 1, 2, \ldots, n$) of this equation there corresponds a system of functions $k_{ij}(x, y)$ (for $j = 1, \ldots, n$) that are determined from the λ_i just as the functions $k_{1j}(x, y)$ (for $j = 1, \ldots, n$) were determined from λ_1. The functions $k_{ij}(x, y)$ have as many continuous derivatives as do the $a_{ls}(x, y)$. Here,

$$\frac{\partial z_i}{\partial x} = \lambda_i \frac{\partial z_i}{\partial y} + f_i^*(x, y, u_1, \ldots, u_n) \qquad (i = 1, 2, \ldots, n),$$

where

$$z_i = \sum_{j=1}^{n} k_{ij} u_j \qquad (i = 1, 2, \ldots, n).$$

It remains to show that $|k_{ij}| \neq 0$. Let us suppose the opposite, that is, that $|k_{ij}(x^0, y^0)| = 0$ at some point (x^0, y^0) of that region in which all the $k_{ij}(x, y)$ are defined. Then, there exist constants C_s, not all equal to zero, such that

$$\sum_s C_s k_{si}(x^0, y^0) = 0 \qquad (i = 1, 2, \ldots, n). \tag{7.9}$$

If we multiply the i th of these equations by a_{ij} and sum over i, we obtain

$$0 = \sum_{i,s} C_s k_{si}(x^0, y^0)\, a_{ij}(x^0, y^0)$$

† We note that for a system consisting of $(n - 1)$ equations (which, as in Section 44 of my book on ordinary differential equations, we shall need to write in canonical form), the assumptions written in italics are valid, so that, by the inductive hypothesis, such a system of $(n - 1)$ equations can be written in canonical form. This is easily proved by expressing the matrix $||a_{ij}|| - \lambda E$ in terms of the corresponding matrix of the transformed system, which is analogous to (134*) of Section 44 of the book referred to.

$$= \sum_s C_s \sum_i k_{si}(x^0, y^0)\, a_{ij}(x^0, y^0) = \sum_s C_s \lambda_s (x^0, y^0)\, k_{sj}(x^0, y^0).$$

We made this last transformation by using the relationship

$$\lambda_s k_{sj} = \sum k_{si} a_{ij},$$

which is analogous to (7.6).

Thus, we have obtained equations analogous to (7.9), where, instead of C_s we have $C_s\lambda_s(x^0, y^0)$. Analogously, we obtain

$$\sum_s C_s \lambda_s^m (x^0, y^0)\, k_{si}(x^0, y^0) = 0 \quad \text{for} \quad m = 2, 3, \ldots, n-1.$$

Since the determinant of the coefficients of the $C_s k_{si}(x^0, y^0)$ in these equations (the van der Monde determinant) is different from zero for the various $\lambda_1, \ldots, \lambda_n$, we see that, for all s and i,

$$C_s k_{si}(x^0, y^0) = 0,$$

which is impossible.

Remarks: (1) It is easy to see that all the reasoning that we used in this section will remain valid for the case in which the coefficients a_{ij} and f_i are complex-valued functions. However, in what follows, we shall assume that a_{ij} and f_i are real functions.

(2) If equation (7.7) has only distinct real roots in the region G in question, it follows from the remarks above that, in a neighbourhood of the point A, to each root λ_i there corresponds a unique solution (precisely determined except for sign) $k_{i1}, k_{i2}, \ldots, k_{in}$ of the system (7.6) such that in that neighbourhood $\sum_{j=1}^{n} k_{ij}^2 = 1$ and the functions k_{ij} have the same smoothness as do the a_{ts} (see the footnote following (7.8)). By using this fact, we may show that, under the conditions stated for the entire region G, if this region is simply connected, there exists a nonsingular linear transformation of the unknown functions that reduces the system (7.1) to the canonical form (7.2). Here, the coefficents of this transformation have the same smoothness as do the a_{ij}, and the system (7.2) takes the form

$$\frac{\partial v_i}{\partial x} = \lambda_i \frac{\partial v_i}{\partial y} + f_i^*(x, y, v_1, \ldots, v_n)$$
$$(i = 1, 2, \ldots, n). \tag{7.10}$$

(3) If equation (7.7) has no real roots anywhere in the entire region G in the xy-plane, the system (7.1) is said to be elliptic in that region.

If throughout the region G there exists a linear nonsingular transformation of the unknown functions u_i with real coefficients possessing the same smoothness as do the $a_{ij}(x, y)$ that reduces the system (7.1) to the form (7.10), the system (7.1) is said to be hyperbolic in the region G.

On the other hand, if, throughout the region G, all the roots λ of equation (7.7) are real and distinct, the system (7.1) is said to be hyperbolic in the narrow sense. It follows from the preceding remark that a system that is hyperbolic in the narrow sense in a simply connected region G is hyperbolic in that region.

In just the same way, the general linear system of partial differential equations with two independent variables

$$\frac{\partial^{n_i} u_i}{\partial t^{n_i}} = \sum_{j=0}^{N} \sum_{k=0}^{n_j - 1} A_{ij}^k(t, x) \frac{\partial^{n_j} u_j}{\partial t^k \, \partial x^{n_j - k}} + \cdots \tag{7.11}$$
$$(i = 1, \ldots, N)$$

is said to be elliptic in the region G if throughout this region, the determinant of the matrix

$$\left\| \begin{matrix} \lambda^{n_1} & & \\ & \ddots & \\ & & \lambda^{n_N} \end{matrix} \right\| - \left\| \sum_{k=0}^{n_j - 1} A_{ij}^k \lambda^k \right\|$$

has no real roots λ.

On the other hand, if all the roots of this determinant are real and distinct, the system (7.11) is said to be hyperbolic in the narrow sense.

Problem. Show that if the system (7.11) is hyperbolic in the narrow sense in a simply connected region G, then a system of first-order equations constructed from equations (7.11) in the same way as the system (2.5) was constructed from equation (2.3) is hyperbolic in that region.

(4) If all the roots of equation (7.7) are real, the transformed system (7.2) can be made real. To do this, we need

to choose real coefficients of the linear transformation from the functions u_i to the functions v_i, which is always possible in this case.

On the other hand, if equation (7.7) possesses complex roots, these roots come in pairs of complex conjugate roots. Then, the system (7.2) can be constructed in such a way that to every equation of the form

$$\frac{\partial v_k}{\partial x} = \lambda_k \frac{\partial v_k}{\partial y} + f_k(x, y, v_1, \ldots, v_n)$$

in this system there will correspond a complex conjugate equation, that is, an equation of the form

$$\frac{\partial v_l}{\partial x} = \lambda_l \frac{\partial v_l}{\partial y} + f_l(x, y, v_1, \ldots, v_n),$$

where

$$v_l = \bar{v}_k;\ \lambda_l = \bar{\lambda}_k;\ f_l(x, y, v_1, \ldots, v_n) = \overline{f_k(x, y, v_1, \ldots, v_n)}.$$

If we separate the real and imaginary parts in these equations and set

$$v_k = w_k^* + i w_k^{**},$$
$$\lambda_k = a_k + i b_k,$$
$$f_k = f_k^* + i f_k^{**},$$

we obtain

$$\frac{\partial w_k^*}{\partial x} = a_k \frac{\partial w_k^*}{\partial y} - b_k \frac{\partial w_k^{**}}{\partial y} + f_k^*,$$
$$\frac{\partial w_k^{**}}{\partial x} = b_k \frac{\partial w_k^*}{\partial y} + a_k \frac{\partial w_k^{**}}{\partial y} + f_k^{**}.$$

The simplest system of this form is the familiar system of Cauchy-Riemann equations

$$\frac{\partial w_1}{\partial x} = -\frac{\partial w_2}{\partial y},$$
$$\frac{\partial w_2}{\partial x} = \frac{\partial w_1}{\partial y}.$$

Analogously, equations of the form

$$\frac{\partial v_k}{\partial x} = \alpha \frac{\partial v_{k-1}}{\partial y} + \lambda \frac{\partial v_k}{\partial y} + f_k(x, y, v_1, \ldots, v_n).$$

can be broken into real and imaginary parts. Thus, we have shown that the system (7.1) can be reduced by a linear nonsingular transformation with smooth real coefficients to canonical form in which all the equations, in contrast with the equations (7.2), are necessarily real. (See the remark to Section 47 of my book on ordinary differential equations referred to above.)

Remark: (5) Consider the system of equations

$$\frac{\partial u_i}{\partial x} = \sum_{j=1}^{n} a_{ij}(x, y, u_1, \ldots, u_n)\frac{\partial u_j}{\partial y} + f_i(x, y, u_1, \ldots, u_n) \quad (i = 1, \ldots, n). \tag{7.12}$$

which is hyperbolic in the narrow sense and quasilinear. For such a system, the quantities λ and k_i that appear in equations (7.6) and (7.7) depend not only on x and y but also on $u_1, \ldots, u_n$. We assume that all the roots of equation (7.7) are real and distinct in some region of variation of the variables $x, y, u_1, \ldots, u_n$.

Suppose that $k_{j1}, \ldots, k_{jn}$ is a nontrivial solution of the system (7.6) for $\lambda = \lambda_j$ (where $j = 1, \ldots, n$). If we multiply the ith equation in the system (7.12) by k_{ji} and sum over all i, we obtain

$$\sum_{i=1}^{n} k_{ji}\left(\frac{\partial u_i}{\partial x} - \lambda_j \frac{\partial u_i}{\partial y}\right) = \tilde{f}_j(x, y, u_1, \ldots, u_n) \quad (j = 1, \ldots, n). \tag{7.13}$$

Each of the equations of the system (7.13) contains derivatives of all the unknown functions in only one direction.

In the case of $n = 2$, the system (7.13) can be reduced to a form analogous to (7.10). Let us denote by $\mu_j(x, y, u_1, u_2)$ a particular solution of the equation

$$\frac{\partial(k_{j1}\mu_j)}{\partial u_2} = \frac{\partial(k_{j2}\mu_j)}{\partial u_1} \quad (j = 1, 2) \tag{7.14}$$

and let us introduce, instead of the functions u_j, new unknown functions $v_j(x, y, u_1, u_2)$ such that

$$\frac{\partial v_j}{\partial u_i} = \mu_j k_{ji} \quad (i, j = 1, 2). \tag{7.15}$$

The relations (7.15) are not contradictory since the μ_j satisfy equations (7.14). If we multiply the j th equation of the system (7.13) by μ_j (for $j=1, 2$), we arrive at the following canonical system of equations for v_1 and v_2:

$$\frac{\partial v_j}{\partial x}=\lambda_j \frac{\partial v_j}{\partial y}+f_j^*(x, y, v_1, v_2) \quad (j=1, 2). \tag{7.16}$$

The functions v_j are often called the generalised Riemann invariants.

For $n>2$, reduction of the system (7.12) to the form (7.16) is in general impossible.

Chapter 2

Hyperbolic equations

PART 1

THE CAUCHY PROBLEM IN THE CLASS OF NONANALYTIC FUNCTIONS

8. CORRECT POSING OF THE CAUCHY PROBLEM

The Kovalevsky theorem asserts the existence of an analytic solution to the Cauchy problem for analytic equations with analytic initial conditions. Many problems in physics are reduced to the Cauchy problem for analytic equations with initial conditions that are several times differentiable but that are not analytic. At first glance, such a method of solving this problem seems natural. We approximate the given initial functions and their derivatives with polynomials. According to a theorem of Weierstrass, polynomials can be chosen such that the difference between them and the corresponding given functions will be arbitrarily small throughout the entire portion of the plane $t = t_0$ on which the Cauchy conditions are given. According to the Kovalevsky theorem for analytic equations, it is possible to solve the Cauchy problem if we replace the original initial conditions with new ones that approximate to the original ones. It would seem natural to expect that this solution of the new Cauchy problem with initial conditions in the form of approximating polynomials will be close to the solution of the same problem with the original initial conditions, at least near that portion of the plane $t = t_0$ on which the Cauchy conditions are given. However, Hadamard exhibited an example that shows that such is not always the case.

Let us consider the following problem. Find the solution to Laplace's equation

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = 0, \tag{8.1}$$

such that, at $t = 0$,

$$u(0, x) = 0, \qquad (8.2)_1$$

$$u_t(0, x) = \frac{1}{n^k} \sin nx, \qquad (8.2)_2$$

where n and k are positive constants. It is easy to see that one solution of this problem is

$$u(t, x) = \frac{1}{n^{k+1}} \sinh nt \sin nx. \tag{8.3}$$

Since

$$|u_t'(0, x)| \leqslant \frac{1}{n^k},$$

for sufficiently large n, the absolute value of $u_t'(0, x)$ will be everywhere arbitrarily small. On the other hand, the solution $u(t, x)$ of the Cauchy problem that we are considering will, as formula (8.3) shows, take arbitrarily large values for arbitrarily small t if n is sufficiently great. The situation remains unchanged if we require not only that $|u_t'(0, x)|$ be everywhere small but also that all the derivatives of $u_t'(0, x)$ with respect to x up to order $k - 1$ also be small. Here, k is an arbitrary positive integer greater than 1. We do not speak of the smallness of the initial values of the function u itself since, from condition $(8.2)_1$ they are everywhere equal to zero.

Let us suppose that we have found the solution to the Cauchy problem for equation (8.1) under certain initial conditions

$$u(0, x) = \varphi_0(x),$$
$$u_t(0, x) = \varphi_1(x).$$

Suppose that this solution is the function $u_0(t, x)$. Then, for the initial conditions

$$u(0, x) = \varphi_0(x); \quad u_t'(0, x) = \varphi_1(x) + \frac{1}{n^k} \sin nx$$

the solution to the Cauchy problem will be the function

$$u_0(t, x) + \frac{1}{n^{k+1}} \sinh nt \sin nx.$$

Thus, a very small change in the initial functions and their first $k-1$ derivatives that is made by adding to the original initial conditions the functions $(8.2)_1$ and $(8.2)_2$ may entail arbitrarily great changes of the form (8.3) in the solution to the Cauchy problem even at points arbitrarily close to the initial value $t=0$.

We shall say that the Cauchy problem is correctly stated in some closed region $\bar{G}$ of the space $t, x_1, \ldots, x_n$ contained in the region G_0 of the hyperplane $t=t_0$ on which the Cauchy conditions are given for a system of linear equations of the form

$$\frac{\partial^{n_i} u_i}{\partial t^{n_i}} = \sum_{j=1}^{N} \sum_{k_0, k_1, \ldots, k_n} A_{ij}^{(k_0, k_1, \ldots, k_n)}(t, x_1, \ldots, x_n) \times \frac{\partial^k u_j}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}} + f_i(t, x_1, \ldots, x_n), \quad (8.4)$$

$$i=1, 2, \ldots, N, \quad k_0+k_1+\ldots+k_n=k \leqslant n_j, \quad k_0 < n_j$$

if there exist positive integers L_1 and L_2 with the following properties:

(1) For arbitrarily continuous functions $\varphi_i^{(k)}(x_1, \ldots, x_n)$, with continuous derivatives up to order L_1 defined in G_0, there exists a unique solution to the system (8.4) in the region $\bar{G}$ such that

$$\frac{\partial^k u_i}{\partial t^k} = \varphi_i^{(k)}(x_1, x_2, \ldots, x_n) \qquad (k=0, 1, \ldots, n_i-1) \quad (8.5)$$

at $t=t_0$;

(2) For an arbitrary positive number ε, a positive number η exists, such that throughout the region $\bar{G}$, the solution to the Cauchy problem changes by an amount less than ε if the functions $\varphi_i^{(k)}$ and their first L_2 derivatives with respect to $x_1, \ldots, x_n$ change by an amount less than η in the region $\bar{G}$.

Ordinarily, the Cauchy conditions are determined from experiments and therefore they cannot be determined with

absolute accuracy. In view of this, in physics (we always use the word 'physics' in a very broad sense) we are interested only in the solutions of the Cauchy problem for those equations for which this problem is correctly stated. As Hadamard's example shows, the Cauchy problem is by no means correctly stated for all equations*.

The considerations made above regarding the correctness of the statement of the Cauchy problem show that there are other boundary problems for partial differential equations that are of interest in natural science only in the case in which the solution of the equation depends continuously on the boundary conditions - the correctness** of the statement of the problem. Each type of equation has its correctly stated boundary problems.

In almost all cases that we have considered up to now, the formulations of such problems were suggested by physical considerations. In particular, the problems listed in Section 1 are just such correctly stated problems.

In the present chapter, the correctness of the statement of the Cauchy problem will be shown for the wave equation in space with a suitable slope of the plane (the carrier of the initial conditions) and for linear hyperbolic systems of first-order partial differential equations with two independent variables. In accordance with what was said in the condition of the problem given in Remark 3 of Section 7, it follows that the Cauchy problem is correctly stated for general linear hyperbolic (in the narrow sense) systems of the form (7.11) with partial derivatives with respect to two independent variables in a simply connected region.

* It is of interest to note that if we consider solutions to the Cauchy problem for Laplace's equation in the class of functions whose absolute values are bounded by some prestated constant, then small changes in the initial conditions will cause only small changes in the solution. See, for example, LAVRENT'YEV, M.M., *Dokl. Akad. Nauk*, **106**, No. 3, 389-390 (1956).

**In every specific case, the concept of correctness of the statement of the problem should be precisely defined. To define the correctness of the Cauchy problem for nonlinear equations, it is natural to treat as possible initial functions $\phi_i^{(k)}(x_1, \ldots, x_n)$ only functions that are close to certain specified functions $\tilde{\phi}_i^{(k)}(x_1, \ldots, x_n)$. It may happen that close to one system of functions $\tilde{\phi}_i^{(k)}(x_1, \ldots, x_n)$ the Cauchy problem is correctly stated and close to another system of functions $\tilde{\phi}_i^{(k)}(x_1, \ldots, x_n)$, it is incorrectly stated.

9. CONCEPT OF GENERALISED SOLUTIONS

In the preceding section, we spoke of the correctness of the stating of the Cauchy problem for sufficiently smooth initial conditions. However, physical problems do not by any means always lead to initial conditions that are sufficiently smooth for it to be possible to assert the existence of a solution to the corresponding problem. If the initial conditions are not continuous and sufficiently many times differentiable, a differentiable solution to the corresponding boundary-value problem may often fail to exist. In this case, the use of so-called 'generalised solutions' of the differential equations turns out to be extremely useful.

The theory of generalised solutions to partial differential equations was worked out by S.L. Sobolev in the thirties. Such solutions are defined either as the limit of a sequence of ordinary solutions or they are defined by means of integral identities.

As an example, let us consider the Cauchy problem for the equation

$$\frac{\partial u}{\partial t}=\frac{\partial u}{\partial x} \tag{9.1}$$

with the initial condition

$$u(0,\ x)=\varphi(x), \tag{9.2}$$

where $\varphi(x)$ is a continuously differentiable function defined on the interval $a \leqslant x \leqslant b$. As can be easily shown, the solution to this problem in the region $D\{a < x+t < b\}$ is the function

$$u(t,\ x)=\varphi(x+t). \tag{9.3}$$

Now suppose that the function $\varphi(x)$ is continuous but not differentiable on the interval $[a,\ b]$. We know that such a function can be represented as the limit of a sequence of continuously differentiable functions $\varphi^{(k)}(x)$ that converges uniformly on the interval $[a,\ b]$. Then, the corresponding solutions $\varphi^{(k)}(x+t)$ of equation (9.1) will converge uniformly in D to the function (9.3). This provides a basis for the function (9.3), which may also be considered a solution of equation (9.1) in the generalised sense.

Definition 1. A system of functions $(u_1,\ \ldots,\ u_N)$ is said to be a generalised solution of a system of differential equations in a region G if there exists an infinite sequence of solutions

$(u_1^{(k)}, \ldots, u_N^{(k)})$ of this system that converges uniformly to $(u_1, \ldots, u_N)$ that is, if

$$\sup_{P \in G} \sum_{i=1}^{N} |u_i(P) - u_i^{(k)}(P)| \to 0 \quad \text{as} \quad k \to \infty.$$

Remark: Sometimes a system of functions $(u_1, \ldots, u_N)$ is also said to be a generalised solution of a system of differential equations in the case in which a sequence of ordinary solutions $(u_1^{(k)}, \ldots, u_N^{(k)})$ converges in mean to $(u_1, \ldots, u_N)$, that is, if

$$\int_G \sum_{i=1}^{N} [u_i(P) - u_i^{(k)}(P)]^2 \, dP \to 0 \qquad \text{as} \quad k \to \infty.$$

Generalised solutions defined in this way may be discontinuous. [See S.L. Sobolev, Uravneniya Matematicheskoi Fiziki (Equations of mathematical physics). Gostekhizdat, 1954 (especially pp. 314, 322, and 329) and Nekotorye Primeneniya funktsional'nogo analiza v matematicheskoy fizike (Certain applications of functional analysis in mathematical physics) Leningrad, 1950.]

Broadening of the class of solutions of any boundary-value problem is of interest in the case in which the uniqueness of the solution is maintained by such broadening. For the most typical boundary-value problems associated with partial differential equations, Sobolev proved the existence and uniqueness of their generalised solutions. In this connection, it is especially important to make it clear how one should understand the boundary-value problems for the generalised solutions.

For linear homogeneous elliptic and parabolic equations with sufficiently smooth coefficients, the introduction of generalised solutions by the procedure shown above does not broaden the class of ordinary solutions to these equations (see Theorem 4 of Section 30). On the other hand, for hyperbolic equations, the class is broadened in a very significant way, as is shown by the very simple example that we considered above.

The introduction of generalised solutions is convenient in that, for the existence of ordinary solutions of the basic boundary-value problems, we sometimes need to impose very stringent smoothness conditions on the functions de-

fined on the boundary of the region in question whereas, for the existence of generalised solutions, such smoothness of the functions given on the boundary is not required. Thus, the generalised solution of the Cauchy problem (9.1)-(9.2) exists, as we have already seen, for an arbitrary continuous function $\varphi(x)$.

Consideration of the generalised solutions of equation (9.1) is all the more natural since the function $\varphi(x)$ itself is usually known to us in an approximate form only. Therefore, the corresponding function $u(t, x)$ given by formula (9.3) is also only an approximation of the exact solution of the problem posed. It is completely immaterial to us whether this approximation is an ordinary or only a generalised solution of equation (9.1). It is important that this solution differ only slightly from the true solution when the function $\varphi(x)$ bears a uniform but slight difference from the actual initial value of $u(0, x)$.

Another way of introducing generalised solutions, also discovered by Sobolev, consists in using integral identities that, for ordinary solutions, are consequences of the equations in question. We shall illustrate this method of introducing generalised functions, which is widespread at the present time, by considering an example of a first-order linear equation.

Suppose that the function $u(x_1, \ldots, x_n)$ is continuously differentiable in the region D and that it satisfies the equation

$$L(u) \equiv \frac{\partial u}{\partial x_1} + \sum_{i=2}^{n} a_i(x_1, \ldots, x_n) \frac{\partial u}{\partial x_i} + b(x_1, \ldots, x_n) u = f(x_1, \ldots, x_n), \tag{9.4}$$

where the a_i are continuously differentiable and b and f are continuous in D. Let us multiply both sides of equation (9.4) by the function $\sigma(x_1, \ldots, x_n)$, which is continuously differentiable in the region D and which vanishes in a neighbourhood of the boundary of D. Let us integrate the resulting equation over the region D. Integrating by parts, we obtain

$$\int\!\!\int_D [uM(\sigma) - f\sigma]\, dx_1 \ldots dx_n = 0, \tag{9.5}$$

where

$$M(\sigma) \equiv -\frac{\partial \sigma}{\partial x_1} - \sum_{i=2}^{n} \frac{\partial (a_i \sigma)}{\partial x_i} + b\sigma.$$

Thus, every ordinary solution to (9.4) satisfies equation (9.5). However, this equation is also satisfied by a wider class of functions $u(x_1, \ldots, x_n)$ since the left side of (9.5) does not contain derivatives of u. Therefore, the following definition is natural.

Definition 2. A function $u(x_1, \ldots, x_n)$ is said to be a generalised solution of equation (9.4) in the region D if equation (9.5) is satisfied for every continuous differentiable function $\sigma(x_1, \ldots, x_n)$ that vanishes at all points of the region D whose distance from the boundary D is less than some positive number ρ_σ (where ρ_σ differs for different σ's).

In considering generalised solutions to boundary-value problems, we need, as was mentioned above, to state in just what sense the boundary conditions are to be understood. Sometimes, these conditions (or some of them) may be considered after changing the integral identity defining the generalised solution. For example, suppose that the region D lies in the half-plane $x_1 \geqslant 0$ and that its boundary Γ contains a segment Γ_1 of the hyperplane $x_1 = 0$.

$$u(0, x_2, \ldots, x_n) = \varphi(x_2, \ldots, x_n) \text{ on } \Gamma_1, \tag{9.6}$$

Then, the piecewise continuous function $u(x_1, \ldots, x_n)$ satisfying the equation for an arbitrary continuously differentiable function $\sigma(x_1, \ldots, x_n)$ that vanishes in a neighbourhood of $\Gamma - \Gamma_1$ is called a generalised solution of the Cauchy problem for equation (9.4) in the region D with initial condition

$$\int\int_D [uM(\sigma) - f\sigma]\, dx_1 \ldots dx_n - \int_{\Gamma_1} \varphi(x_2, \ldots, x_n)\, \sigma(0, x_2, \ldots, x_n)\, dx_2 \ldots dx_n = 0 \tag{9.7}$$

Problem 1. Suppose that a function $u(x_1, \ldots, x_n)$ is continuously differentiable in a closed region $\bar{D}$ and that it is a generalised solution to the Cauchy problem (9.4)-(9.6) in the sense of equation (9.7). Show that this function satisfies equation (9.4) and the initial condition (9.6) in the ordinary sense.

2. Construct a generalised solution (in the sense of equation (9.7)) of the Cauchy problem for equation (9.1) with the initial condition

$$u(0, x) = \begin{cases} -1 & \text{for } x \leqslant 0, \\ 1 & \text{for } x > 0. \end{cases}$$

3. Suppose that a function $u(t, x)$ is a generalised solution of equation (9.4) in the sense of definition 1. Show that this function is also a generalised solution of (9.4) in the sense of definition 2.

10. THE CAUCHY PROBLEM FOR HYPERBOLIC SYSTEMS WITH TWO INDEPENDENT VARIABLES

1. Consider the system*

$$\frac{\partial u_i}{\partial t} - \lambda_i \frac{\partial u_i}{\partial x} = \sum_{j=1}^{N} a_{ij}(t, x) u_j + b_i(t, x) \qquad (i = 1, 2, \ldots, N) \tag{10.1}$$

We shall assume that this system is hyperbolic throughout the entire region in question, that is, that all the λ_i are real functions of t and x. We shall also assume that all the $\lambda_i(t, x)$ are distinct and numbered in increasing order.

Through every point in our region there pass N real characteristics L_i with angular coefficients $k_i = -\frac{1}{\lambda_i}$ in relation to the x-axis (see example 5 of Section 3).

If we do not assume that the coefficients in the system (10.1) are analytic, we cannot draw any conclusions regarding the solvability of the Cauchy problem for this system from the Kovalevsky theorem. We assume that in some closed region $\overline{G}$ bounded by the segment $[a, b]$ of the x-axis and the characteristics L_1 and L_N which pass through the points $(0, a)$ and $(0, b)$ respectively (see Fig. 2), the functions a_{ij}, b_i, and λ_i are continuous and have continuous first derivatives. Let us define N continuously differentiable functions

* Everything stated in the present section can with very slight changes be applied to systems of the form

$$\frac{\partial u_i}{\partial t} - \lambda_i \frac{\partial u_i}{\partial x} = f_i(t, x, u_1, \ldots, u_N) \qquad (i = 1, 2, \ldots, N)$$

under the assumption that the functions $f_i(t, x, u_1, \ldots, u_N)$ have continuous first and second derivatives (see the proof of an existence of a solution to the equation $dy/dx = f(x, y)$ by the method of successive approximations.)

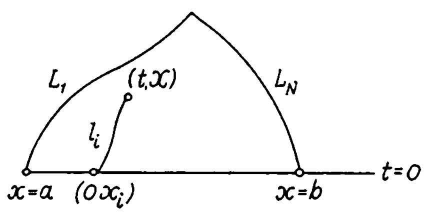
$\varphi_1(x), \ldots, \varphi_N(x)$ on the segment $[a, b]$ and let us pose the Cauchy problem for the system (10.1) in the following manner:

Find a solution $u_1, u_2, \ldots, u_N$ of the system (10.1) that is continuous in $\bar{G}$ and that has continuous first derivatives in G, such that, for $t=0$,

$$u_i(0, x) = \varphi_i(x) \qquad (i = 1, \ldots, N). \tag{10.2}$$

Under the assumptions made, the problem has a unique solution.*

Proof: Let us consider the ith equation of the system (10.1). Its left member is, up to a constant factor, the deriva-



Curves L_1 and L_N do not necessarily intersect. G region may therefore be infinite. G should be bounded by using straight line $t = T$. This also applies for region G situated in half-plane $t < 0$

Fig. 2

tive of the function $u_i(t, x)$ along the curve L_i. Indeed, if we denote by α_i the angle between the tangent to the curve L_i at the point (t, x) and the x-axis, then, as was shown,

* The assumption that all the λ_i are distinct is not essential. All the statements to be made remain valid for the case in which some of the λ_i coincide. In such a case, instead of the characteristic L_1, which originates at the point (0, 2), we would need to take for the definition of G a solution of the equation $\frac{dx}{dt} = -\lambda_{\min}(t, x)$, passing through the point $(0, a)$, where

$$\lambda_{\min}(t, x) = \min\{\lambda_1(t, x), \ldots \lambda_N(t, x)\},$$

and, instead of L_N, a solution of the equation $\frac{dx}{dt} = -\lambda_{\max}(t, x)$, passing through the point $(0, b)$, where

$$\lambda_{\max}(t,x) = \max\{\lambda_1(t, x), \ldots, \lambda_N(t, x)\}.$$

The functions $\lambda_{\min}(t, x)$ and $\lambda_{\max}(t, x)$ are continuous and, as is easily shown, they satisfy a Lipschitz condition with respect to t if all the λ_i are continuous and have bounded derivatives with respect to t.

$$\tan\alpha_i = -\frac{1}{\lambda_i}.$$

Consequently,

$$\cos\alpha_i = -\frac{\lambda_i}{\sqrt{1+\lambda_i^2}}; \qquad \sin\alpha_i = \frac{1}{\sqrt{1+\lambda_i^2}},$$

and

$$\frac{\partial u_i}{\partial s_i} = \left(\frac{\partial u_i}{\partial t} - \lambda_i\frac{\partial u_i}{\partial x}\right)\frac{1}{\sqrt{1+\lambda_i^2}}.$$

Here, s_i denotes the arc length of the characteristic L_i, and $\frac{\partial}{\partial s_i}$ denotes differentiation in the direction of the characteristic L_i.

The system (10.1) may be rewritten in the form

$$\sqrt{1+\lambda_i^2}\frac{\partial u_i}{\partial s_i} = \sum_{i=1}^{N} a_{ij}u_j + b_i \quad (i = 1, 2, \ldots, N). \tag{10.3}$$

If we denote by $\underset{i}{du_i}$ the differential of the function u_i as we move along the curve L_i, we obtain, from equation (10.3),

$$\underset{i}{du_i} = (\sum a_{ij}u_j + b_i)\frac{ds_i}{\sqrt{1+\lambda_i^2}}$$

and since $ds_i = \sqrt{1+\lambda_i^2}dt$ we have

$$\underset{i}{du_i} = (\sum a_{ij}u_j + b_i)\,dt \qquad (i = 1, 2, \ldots, N). \tag{10.4}$$

Let us now fix an arbitrary point (t, x) in the region $\bar{G}$ and let us denote by l_i that portion of the corresponding curve L_i between the point (t, x) and its intersection, at a point $(0, x_i)$ with the segment $[a, b]$ of the $(t = 0)$-axis (see Fig. 2). We then integrate the ith of equations (10.4) along the arc l_i from the point $(0, x_i)$ to the point (t, x). We then obtain the system of integral equations

$$u_i(t, x) - u_i(0, x_i) = \int_{l_i}\left(\sum_{j=1}^{N} a_{ij}u_j + b_i\right)dt.$$

$$(i = 1, 2, \ldots, N),$$

or, on the basis of the initial conditions (10.2),

$$u_i(t, x) = \varphi_i(x_i) + \int\limits_{l_i} \left(\sum_{j=1}^{N} a_{ij} u_j + b_i \right) dt. \tag{10.5}$$

Obviously, every solution to the system (10.1) that satisfies the initial conditions (10.2) is a solution of the system (10.5). Conversely, if we have a solution of the system of integral equations (10.5) and if the functions constituting this solution have continuous derivatives with respect to t and x in G , then, by carrying out operations inverse to those by means of which we arrived at (10.5) from (10.1), we see that the solution to the system (10.5) is also a solution of the posed Cauchy problem. Thus, the problem is reduced to proving the existence of a continuously differentiable solution to the system (10.5).

Let us construct successive approximations to the solution of the system (10.5) in the following manner: We set

$$u_i^{(0)}(t, x) = \varphi_i(x_i) \qquad (i = 1, \ldots, N),$$

$$u_i^{(1)}(t, x) = \varphi_i(x_i) + \int\limits_{l_i} \left[\sum_{j=1}^{N} a_{ij}(t, x) u_j^{(0)} + b_i(t, x) \right] dt$$

$$(i = 1, 2, \ldots, N)$$

and, in general,

$$u_i^{(n+1)}(t, x) = \varphi_i(x_i) + \int\limits_{l_i} \left[\sum_{j=1}^{N} a_{ij}(t, x) u_j^{(n)} + b_i(t, x) \right] dt$$

$$(i = 1, \ldots, N).$$

The last equation should more properly be written in the form

$$\begin{aligned} u_i^{(n+1)}(t, x) &= \varphi_i[x_i(0, t, x)] \\ &+ \int\limits_0^t \Big[\sum_{j=1}^{N} a_{ij}(\tau, x_i(\tau, t, x)) u_j^{(n)}(\tau, x_i(\tau, t, x)) \\ &+ b_i(\tau, x_i(\tau, t, x)) \Big] d\tau \qquad (i = 1, 2, \ldots, N). \end{aligned}$$

We assume that $x = x_i(t, t^0, x^0)$ is the equation of the characteristic L_i passing through the point (t^0, x^0). If we can show the uniform convergence of the sequences $u_i^{(n)}(t, x)$ in

a closed region $\bar{G}$, the system of limiting functions $u_i(t, x)$ will satisfy equations (10.5). The uniform convergence of the sequence $u_i^{(n)}(t, x)$ is equivalent to the uniform convergence of the series

$$u_i^{(0)}(t, x) + \sum_{n=0}^{\infty} [u_i^{(n+1)}(t, x) - u_i^{(n)}(t, x)]. \tag{10.6}$$

To show that this series converges uniformly, let us construct a numerical majorant for it. Since the functions $u_i^{(0)}(t, x)$ and $u_i^{(1)}(t, x)$ are continuous in the closed region G, they are bounded in that region. Let us therefore set

$$M = \max \{ |u_1^{(0)}|, \ldots, |u_N^{(0)}|, |u_1^{(1)}|, \ldots, |u_N^{(1)}| \}$$

in the region $\bar{G}$. Then,

$$|u_i^{(0)}(t, x)| \leqslant M,$$
$$|u_i^{(1)} - u_i^{(0)}| \leqslant 2M,$$
$$(t, x) \in \bar{G}.$$

Let us denote $\max |a_{ij}|$ in the region $\bar{G}$ for all $i, j = 1, \ldots, N$ by A. Then,

$$|u_i^{(2)}(t, x) - u_i^{(1)}(t, x)| \leqslant \int_{l_i} \sum_{j=1}^{N} |a_{ij}| \cdot |u_j^{(1)} - u_j^{(0)}| dt \leqslant 2MANt,$$

$$|u_i^{(3)}(t, x) - u_i^{(2)}(t, x)|$$
$$\leqslant \int_{l_i} \sum_{j=1}^{N} |a_{ij}| \cdot |u_j^{(2)} - u_j^{(1)}| dt \leqslant 2MA^2N^2 \frac{t^2}{2}.$$

Let us now assume that

$$|u_i^{(n)}(t, x) - u_i^{(n-1)}(t, x)| \leqslant 2M \frac{A^{n-1}N^{n-1}t^{n-1}}{(n-1)!}.$$

Then,

$$|u_i^{(n+1)}(t, x) - u_i^{(n)}(t, x)|$$
$$\leqslant \int_{l_i} \sum_{j=1}^{N} |a_{ij}| \cdot |u_j^{(n)} - u_j^{(n-1)}| dt \leqslant 2M \frac{A^nN^nt^n}{n!}.$$

Thus, by mathematical induction, we have, for arbitrary n,

$$|u_i^{(n+1)}(t, x) - u_i^{(n)}(t, x)| \leqslant 2M\frac{(ANt)^n}{n!}.$$

But the region G is bounded and, by taking a fixed number T exceeding all values of t in this region, we see that throughout the region G

$$|u_i^{(n+1)} - u_i^{(n)}| \leqslant 2M\frac{(ANT)^n}{n!}.$$

Since the numerical series

$$\sum\frac{(ANT)^n}{n!}$$

converges, it follows that the series (10.6) converges uniformly in the entire region $\overline{G}$, which in turn proves the existence of a solution to the system (10.5) and its continuity.

Let us now prove the uniqueness of a solution to the system (10.5) that is continuous, and hence bounded, in $\overline{G}$. Suppose that we have two such solutions to the system (10.5), namely, $u_1, \ldots, u_N$ and $\breve{u}_1, \ldots, \breve{u}_N$. If we substitute both these solutions into the system and subtract the corresponding equations one from the other, we obtain

$$u_i(t, x) - \breve{u}_i(t, x) = \int_{l_i} \sum_{j=1}^{N} a_{ij}(u_j - \breve{u}_j)\, dt.$$

Let us now suppose that

$$\max_{\substack{(x, t)\in\overline{G} \\ i=1, \ldots, N}} |u_i - \breve{u}_i| = M > 0.$$

Then by making successive estimates of the difference $|u_i(t, x) - \breve{u}_i(t, x)|$, as was done in the proof of the existence, we obtain

$$M \leqslant M\frac{(ANT)^n}{n!}$$

for any arbitrary n, which leads to a contradiction as soon as n is sufficiently large. Consequently, $M = 0$ and

$$u_i(t, x) = \breve{u}_i(t, x) \quad (i = 1, 2, \ldots, N),$$

which means that the solution is unique.

To complete the proof, we still need to show that these functions $u_i(t, x)$ have continuous first derivatives with respect to t and x. Obviously, it will be sufficient to show that the functions $u_i(t, x)$ have continuous first derivatives in the direction of l_i and continuous first derivatives with respect to x at every point since this and the smoothness of l_i will imply continuity of the derivatives with respect to t and x throughout the region G.

The existence and continuity of the derivatives $u_i(t, x)$ along l_i follow directly from the system (10.5) and from the continuity of the solution obtained. To prove the existence and continuity of the derivatives $\frac{\partial u_i}{\partial x}$, let us note first that the assumed continuous differentiability of $\varphi_i(x)$, $\lambda_i(t, x)$, $a_{ij}(t, x)$, and $b_i(t, x)$ implies that all the approximations made in the proof of the existence of the solution have continuous derivatives with respect to x. Let us differentiate with respect to x the equation defining the $(n+1)$st approximation. We obtain*

$$\frac{\partial u_i^{(n+1)}(t, x)}{\partial x} = \varphi_i'[x_i(0, t, x)] \frac{\partial x_i(0, t, x)}{\partial x}$$
$$+\int_0^t \Bigg[\sum_j \frac{\partial a_{ij}(\tau, x_i(\tau, t, x))}{\partial x} u_j^{(n)}(\tau, x_i(\tau, t, x))$$
$$+\sum_j a_{ij}(\tau, x_i(\tau, t, x)) \frac{\partial u_j^{(n)}(\tau, x_i(\tau, t, x))}{\partial x_i} \frac{\partial x_i(\tau, t, x)}{\partial x}$$
$$+\frac{\partial b_i(\tau, x_i(\tau, t, x))}{\partial x}\Bigg] d\tau$$
$$(i = 1, 2, \ldots, N).$$

From the assumptions made with regard to the system (10.1), we can show that the sequence of derivatives $\frac{\partial u_i^{(n)}}{\partial x}$ (for $n = 1, 2, \ldots$) converges uniformly by the exact same method as that used to show the convergence of the sequence of $u_i^{(n)}$; only the constants change in the estimates. Therefore,

* The coordinate x_i of the point of the intersection l_i with the straight line $t = \tau$ is a continuously differentiable function of t and x because of the assumed continuity of the derivatives with respect to λ_i. The limits of integration with respect to t in the line integral do not change with changing x.

By using inequality (10.9) and again making an estimate $|z_i(t, x)|$ by means of equation (10.8), we obtain

$$|z_i(t, x)| \leqslant \eta(1 + ANt) + \varepsilon \frac{A^2N^2t^2}{2!}.$$

By repeating this operation n times, we can prove the inequality

$$|z_i(t, x)| \leqslant \eta\left(1 + ANt + \ldots + \frac{(ANt)^{n-1}}{(n-1)!}\right) + \varepsilon \frac{(ANt)^n}{n!}.$$

Taking the limit as $n \to \infty$, we obtain

$$\varepsilon \leqslant \eta e^{ANT}.$$

From this it is clear that $\varepsilon \to 0$ as $\eta \to 0$ since e^{ANT} is a constant not depending on η.

Problem 1. Formulate a definition of a generalised solution to the Cauchy problem for the system (10.1) under conditions (10.2) in a manner analogous to what was done in Section 9 for equation (9.4).

2. Prove the uniqueness of the generalised solution of the Cauchy problem (10.1)-(10.2) in the class of continuously differentiable functions outside a finite number of smooth curves.

3. Suppose that the generalised solution of the Cauchy problem (10.1)-(10.2) has a finite number of smooth curves of disconitnuity of the first kind outside of which it ix continuously differentiable. Show that these curves are characteristics of the system (10.1).

3. We shall conclude the present section with a brief description of the method of finite differences, which is convenient for obtaining an approximate solution of the Cauchy problem stated in subsection 1.

Suppose that we are given initial functions $\varphi_i(x)$ on the interval $[a, b]$ of the x-axis. To find approximate values of the function $u_i(t, x)$ satisfying the system (10.1) and taking at $t = 0$ the given values of $\varphi_i(x)$, we proceed as follows: Let us take some integer n and let us partition the interval $[a, b]$ into n equal portions of length $h = \frac{b-a}{n}$. We then draw straight lines $x = a + ph$ and $t = qh$ for those integral values of p and q such that the region G in which a

solution to the Cauchy problem (see subsection 1 of the preceding section) will be covered with a square net each square in which is of side h. Let us number the corners of the squares with two indices; specifically, let us denote by M_{pq} the point of intersection of the two straight lines $x=a+ph$ and $t=qh$. We are given the values of the unknown functions $u_i(t, x)$ at all the points M_{p0}:

$$u_i(0, a+ph)=\varphi_i(a+ph)=\varphi_i(M_{p0}).$$

The following is a process by means of which it is possible to find approximate values of $u_i(t, x)$ at all corners of the squares of the net lying within G. At each of the points M_{p0} the coefficients of the system (10.1) are defined; in particular, the N numbers $\lambda_i(M_{p0})=\lambda_i^{p0}$ (for $i=1, \ldots, N$) are defined. From each point M_{p0}, let us draw N segments of straight lines with angular coefficients $k_i^{p0}=-\frac{1}{\lambda_i^{p0}}$ to the intersection with the straight line $t=h$, and let us find the values of $u_i(t, x)$ at the opposite ends of the corresponding segments. To do this, we use the form (10.4) of the system (10.1) and we replace the differential (as we move along the characteristic L_i) with an increment and the corresponding exact equation with an approximate one. We then obtain the relationship

$$\underset{i}{\Delta} u_i \approx \left(\sum a_{ij}u_j+b_i\right) h,$$

which makes it possible to find the increment in the function $\underset{i}{\Delta} u_i$ that results when we change from the point M_{p0} along the characteristic L_i (more precisely, along the tangent to that characteristic) to the straight line $t=h$.

When we add the increments obtained to the original values of the function at the points M_{p0}, we obtain values of each of the functions u_i at points of the straight line $t=h$. Here, the values of the different functions will generally be determined at different points. By means of some sort of interpolational process carried out on the values found for u_i on the straight line $t=h$, we determine its value at the points M_{p1}, that is, at the corner points of the grid that lie on that straight line. We may then determine the values of $u_i(t, x)$ by the same method at points of the straight line $t=2h$, that lie within the region G. By repeating the inter-

polation and subsequent determination of the values of $u_i(t,x)$ as many times as we may find it necessary, we can find approximate values of all the functions $u_i(t, x)$ at all corners of squares lying in the region G.

It can be shown, that, as $n \to \infty$, the approximate values of the functions converge uniformly to a limit, which represents the exact solution of the Cauchy problem. Consequently, for sufficiently large n, the approximations found by the method described differ arbitrarily slightly from the true solution.

If $N=2$, the process of approximate calculation of the solution to the Cauchy problem is considerably simplified. Then, we have only two families of characteristics. If we partition into small sub-intervals the interval $[a, b]$ and the x-axis on which the initial values of u_1 and u_2 are given and if we draw the tangents to the characteristics of the different families through the points of division to the intersection that is closest to the interval $[a, b]$, we find approximate values of u_1 and u_2 at these points of intersection, as was described above. If we draw tangents to the characteristics through these new points, we thus find approximate values of u_1 and u_2 at those points of intersection of the new tangents that are closest to the interval $[a, b]$, and so forth. Thus, we obtain values of u_1 and u_2 in some sufficiently dense set of points if the initial partition of the interval $[a, b]$ is sufficiently fine. In this case, no square grid or interpolation is required.

11. THE CAUCHY PROBLEM FOR THE WAVE EQUATION. THEOREM ON THE UNIQUENESS OF THE SOLUTION

Theorem. Suppose that a function $u(t, x_1, x_2)$ satisfies the equation

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2} \tag{11.1}$$

within a circular cone K with axis parallel to the t-axis, with vertex at the point A, and with generators that form an angle $\alpha=45°$ with the t-axis. Suppose also that the function $u(t, x_1, x_2)$ and its first and second derivatives are continuous within K and on its boundary.

Then, the value of $u(t, x_1, x_2)$ is uniquely determined

at the point A by the values of u and $\frac{\partial u}{\partial t}$ on the base of the cone lying in the plane $t = t_0$.

The cone K is called the characteristic cone. It is easy to see that the lateral surface of K is a characteristic surface in the sense of subsection 2 of Section 3.

The theorem holds also in the case in which the coordinate $t > t_0$ just as it does in the case in which $t < t_0$.

Remarks: (1) Instead of equation (11.1) we could have taken the equation

$$\frac{\partial^2 u}{\partial t^2} = a^2 \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right), \tag{11.2}$$

where a is any positive constant, in the formulation of the theorem by replacing the cone whose generators form a 45° angle with the t-axis with another cone whose generators form an angle $\alpha =$ arctan a with the t-axis (since equation (11.2) reduces to equation (11.1) when we replace t with t/a.

(2) We can always assume that $t_0 = 0$. The case of any other value of t_0 reduces to this one if we replace the independent variable t with a new independent variable $t^* = t - t_0$. This change of variable does not change the form of equation (11.1).

(3) Suppose that we are given a region G_0 in the plane $t = 0$. Let us construct cones K whose bases lie in the region G_0, whose axes are parallel to the t-axis, and whose generators form an angle of $\pm 45°$ with the t-axis. Then, it follows from our theorem that stating the values of u and $\frac{\partial u}{\partial t}$ in the region G_0 uniquely determines the solution to equation (11.1) in the region G of the space (t, x_1, x_2) filled by the cones K. For example, stating the values of u and $\frac{\partial u}{\partial t}$ in the square $|x_1| < a$, $|x_2| < a$ determines uniquely the twice continuously differentiable solution $u(t, x_1, x_2)$ of equation (11.1) within each of the two pyramids for which this square is the common base and which have lateral surfaces forming a 45° angle with that common base.

(4) Giving the values of u and $\frac{\partial u}{\partial t}$ in some circle G_0 in the

$x_1 x_2$-plane does not determine the solution $u(t, x_1, x_2)$ of equation (11.1) at any point B lying outside the corresponding cones K with common base G_0, with axes parallel to the t-axis and with generators making a 45° angle with the t-axis. To prove this assertion, we need only show that a solution $\tilde{u}(t, x_1, x_2)$ exists such that $\tilde{u}$ and $\frac{\partial \tilde{u}}{\partial t}$ are equal to zero in the circle G_0 and that $\tilde{u}(B) \neq 0$. To construct such a solution, we note that, for an arbitrary twice continuously differentiable function $f(z)$ and for $\alpha_1^2 + \alpha_2^2 = 1$, the function

$$f(t + \alpha_1 x_1 + \alpha_2 x_2) \tag{11.3}$$

is a solution to equation (11.1). (The reader should verify this for nimself.)

The function (11.3) has constant values on every plane

$$t + \alpha_1 x_1 + \alpha_2 x_2 = C, \tag{11.4}$$

each of which makes a 45° angle with the t-axis. Let us choose α_1 and α_2 such that the plane of the family (11.4) that passes through the point B will not intersect the circle G_0. Then, we can choose a twice continuously differentiable function $f(z)$ in such a way that $f(t + \alpha_1 x_1 + \alpha_2 x_2)$ will be nonzero at the point B but zero in G_0. Then,

$$\tilde{u}(t, x_1, x_2) = f(t + \alpha_1 x_1 + \alpha_2 x_2)$$

will be the desired solution.

Remarks: (5) The proof of the uniqueness theorem that we shall give below is applicable for twice continuously differentiable solutions of the equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \ldots + \frac{\partial^2 u}{\partial x_n^2}$$

for arbitrary n. In such a case, the three-dimensional cone K referred to in the formulation of the theorem must be replaced with a cone in the $n+1$-dimensional space with axis parallel to the t-axis and generators forming a 45° angle with the t-axis. This cone is still referred to as the characteristic cone. For $n = 1$, this cone is replaced with a triangle whose base is parallel to the x-axis and whose lateral sides form a 45° angle with it.

Proof of the uniqueness theorem: Suppose that two conti-

nuous solutions $u_1(t, x_1, x_2)$ and $u_2(t, x_1, x_2)$ to equation (11.1) with continuous first and second derivatives are defined within the cone K and on its surface and suppose that u_1 and u_2 and their first derivatives with respect to t coincide on the base of K. Then, the difference

$$u(t, x_1, x_2) = u_2(t, x_1, x_2) - u_1(t, x_1, x_2)$$

must also satisfy the homogeneous equation (11.1) within K, and $u(t, x_1, x_2)$ and $u_t(t, x_1, x_2)$ must vanish on the base of that cone. The uniqueness theorem will be proven if we can show that $u(t, x_1, x_2) = 0$ at the vertex of K. To show this, let us integrate the expression

$$\frac{\partial u}{\partial t}\left(\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2}\right),$$

(which is everywhere equal to 0 since the function u satisfies equation (11.1)). Since

$$\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2} = \frac{1}{2}\frac{\partial}{\partial t}\left(\frac{\partial u}{\partial t}\right)^2,$$

and

$$\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial x_i^2} = \frac{\partial}{\partial x_i}\left(\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_i}\right) - \frac{\partial^2 u}{\partial t\,\partial x_i}\frac{\partial u}{\partial x_i},$$

it follows that

$$0 = \iiint\limits_K \frac{\partial u}{\partial t}\left(\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2}\right) dt\, dx_1\, dx_2$$

$$= \iiint\limits_K \left\{\frac{1}{2}\frac{\partial}{\partial t}\left[\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial u}{\partial x_1}\right)^2 + \left(\frac{\partial u}{\partial x_2}\right)^2\right]\right.$$

$$\left. - \frac{\partial}{\partial x_1}\left(\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_1}\right) - \frac{\partial}{\partial x_2}\left(\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_2}\right)\right\} dt\, dx_1\, dx_2.$$

Let us transform this integral into a double integral by using Ostrogradskii's formula. If we denote the lateral surface of the cone K by K_1 and its base by C, then, since the initial conditions indicate that

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x_1} = \frac{\partial u}{\partial x_2} = 0,$$

there remains * only the integral over K_1.

$$0=\frac{1}{2}\iint_{K_1}\left\{\left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]\cos(n,\ t)\right.$$
$$\left.-2\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_1}\cos(n,\ x_1)-2\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_2}\cos(n,\ x_2)\right\}d\sigma \tag{11.5}$$

But on the lateral surface of the characteristic cone,

$$\cos^2(n,\ t)-\cos^2(n,\ x_1)-\cos^2(n,\ x_2)=0. \tag{11.6}$$

If we multiply and divide the integrand by $\cos(n,\ t)$ and use equation (11.6), we obtain from (11.5),

$$\frac{1}{2\cos(n,\ t)}\iint_{K_1}\left\{\left(\frac{\partial u}{\partial t}\cos(n,\ x_1)-\frac{\partial u}{\partial x_1}\cos(n,\ t)\right)^2\right.$$
$$\left.+\left(\frac{\partial u}{\partial t}\cos(n,\ x_2)-\frac{\partial u}{\partial x_2}\cos(n,\ t)\right)^2\right\}d\sigma=0. \tag{11.7}$$

Here, it was possible to take $\cos(n,\ t)$ from under the integral sign since this quantity is constant on K_1. Specifically,

$$\cos(n,\ t)=\frac{1}{\sqrt{2}}$$

for $t>t_0$ and

$$\cos(n,\ t)=-\frac{1}{\sqrt{2}}$$

for $t<t_0$.

It follows from equation (11.7) that, on the lateral surface of the cone K,

$$\frac{u'_t}{\cos(n,\ t)}=\frac{u'_{x_1}}{\cos(n,\ x_1)}=\frac{u'_{x_2}}{\cos(n,\ x_2)}=v. \tag{11.8}$$

* We call attention to the fact that in the transformations of the integral

$$\iiint_K \frac{\partial u}{\partial t}\left(\frac{\partial^2 u}{\partial t^2}-\frac{\partial^2 u}{\partial x_1^2}-\frac{\partial^2 u}{\partial x_2^2}\right)dt\,dx_1\,dx_2$$

we made use of the continuity of the first derivatives of u within and on the boundary of K and of the integrability over K of the second derivatives of u. The second derivatives of u are always integrable over K if they are continuous within K and on its boundary.

If we denote by m the direction of some generator of the cone K, then, by using equations (11.8), we obtain

$$\frac{\partial u}{\partial m}=u'_t \cos(m,\ t)+u'_{x_1}\cos(m,\ x_1)+u'_{x_2}\cos(m,\ x_2)$$
$$=v\,[\cos(n,\ t)\cos(m,\ t)+\cos(n,\ x_1)\cos(m,\ x_1)$$
$$+\cos(n,\ x_2)\cos(m,\ x_2)]=v\cos(m,\ n)=0\,.$$

(Note that $\cos(m,n)=0$ since the generator of a cone always makes a right angle with the normal to its surface.)

Thus, on the surface of the cone K, the derivative of u in the direction of the generator is equal to zero. It follows from this that the function u is equal to zero at the vertex of the cone since it is equal to zero on its base. This concludes the proof of the uniqueness theorem.

12. FORMULAE RESOLVING THE CAUCHY PROBLEM FOR THE WAVE EQUATION

1. Suppose that two functions $\varphi_0(x_1,\ x_2,\ x_3)$ and $\varphi_1(x_1,\ x_2,\ x_3)$, are defined in some region G_0 of the space $(x_1,\ x_2,\ x_3)$. Suppose also that φ_0 and its first three derivatives are continuous and that φ_1 and its first two derivatives are continuous. We wish to find the solution $u(t,\ x_1,\ x_2,\ x_3)$ of the equation

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}, \tag{12.1}$$

that satisfies at $t=0$ the conditions

$$u(0,\ x_1,\ x_2,\ x_3)=\varphi_0(x_1,\ x_2,\ x_3), \qquad (12.2)_1$$

$$u'_t(0,\ x_1,\ x_2,\ x_3)=\varphi_1(x_1,\ x_2,\ x_3). \qquad (12.2)_2$$

This solution will be defined at all points $(t,\ x_1,\ x_2,\ x_3)$ that are vertices of the characteristic cones whose bases belong to G_0.

Let us first find a solution $u_\varphi(t,\ x_1,\ x_2,\ x_3)$ of equation (12.1) under initial conditions of the particular form

$$u_\varphi(0,\ x_1,\ x_2,\ x_3)=0, \qquad (12.3)_1$$

$$u'_{\varphi t}(0,\ x_1,\ x_2,\ x_3)=\varphi(x_1,\ x_2,\ x_3). \qquad (12.3)_2$$

Then, it is easy to see that the function

$$v(t, x_1, x_2, x_3) = \frac{\partial u_\varphi}{\partial t}$$

satisfies at $t=0$ the conditions

$$v(0, x_1, x_2, x_3) = \varphi(x_1, x_2, x_3),$$

$$v_t'(0, x_1, x_2, x_3) = \frac{\partial^2 u_\varphi}{\partial t^2} = \frac{\partial^2 u_\varphi}{\partial x_1^2} + \frac{\partial^2 u_\varphi}{\partial x_2^2} + \frac{\partial^2 u_\varphi}{\partial x_3^2} = 0.$$

Therefore, if u_φ has continuous third-order derivatives, the solution of equation (12.1) satisfying both conditions (12.2) is given by the formula

$$u = \frac{\partial u_{\varphi_0}}{\partial t} + u_{\varphi_1}. \tag{12.4}$$

Thus, the general Cauchy problem for equation (12.1) is reduced to finding u_φ. We assert that

$$u_\varphi(t, x_1, x_2, x_3) = \frac{1}{4\pi} \iint\limits_{S_t(x_1, x_2, x_3)} \frac{\varphi(\alpha_1, \alpha_2, \alpha_3)}{t}\, d\sigma_t. \tag{12.5}$$

This is known as Kirchhoff's formula. Here, $S_t(x_1, x_2, x_3)$ denotes the sphere of radius t with centre at the point (x_1, x_2, x_3) on the hyperplane $t=0$ on which the function φ is given and $d\sigma_t$ is an element of the surface of that sphere. We shall assume that the function $\varphi(x_1, x_2, x_3)$ and its first k derivatives (for $k \geqslant 2$) are continuous and bounded. Then, as we shall see later from formula (12.6), the function u_φ will also have continuous derivatives up to the kth order inclusively.

Let us first show that the function u_φ given by formula (12.5) satisfies initial conditions (12.3). The first of these conditions is satisfied since

$$\left| \iint\limits_{S_t} \frac{\varphi(\alpha_1, \alpha_2, \alpha_3)}{t}\, d\sigma_t \right| \leqslant \max|\varphi| \cdot \frac{4\pi t^2}{t},$$

and, consequently,

$$u_\varphi(t, x_1, x_2, x_3) \to 0 \text{ as } t \to 0.$$

To check the second condition, we note that if we set

$$\alpha_k = x_k + \beta_k t,$$

we reduce the integral (12.5) to the form

$$u_\varphi(t, x_1, x_2, x_3) = \frac{t}{4\pi} \iint_{S_1} \varphi(x_1 + t\beta_1, x_2 + t\beta_2, x_3 + t\beta_3)\, d\sigma_1, \tag{12.6}$$

where the integration is taken over the sphere S_1 that is fixed for all x_1, x_2, x_3, t :

$$\beta_1^2 + \beta_2^2 + \beta_3^2 = 1, \quad d\sigma_1 = \frac{d\sigma_t}{t^2}.$$

Therefore,

$$\frac{\partial u_\varphi}{\partial t} = \frac{1}{4\pi} \iint_{S_1} \varphi(x_1 + t\beta_1, x_2 + t\beta_2, x_3 + t\beta_3)\, d\sigma_1$$
$$+ \frac{t}{4\pi} \iint_{S_1} \sum_{k=1}^{3} \beta_k \varphi_k (x_1 + t\beta_1, x_2 + t\beta_2, x_3 + t\beta_3)\, d\sigma_1 \tag{12.7}$$

Here, φ_k denotes the derivative of φ with respect to α_k. It is easy to see that the first term on the right side approaches $\varphi(x_1, x_2, x_3)$ as t approaches 0 and that the second term approaches zero since the integral in it remains bounded.

Now, it will be sufficient to show that u_φ defined from Kirchhoff's formula satisfies equation (12.1). From equation (12.6), we obtain

$$\frac{\partial^2 u_\varphi}{\partial x_1^2} + \frac{\partial^2 u_\varphi}{\partial x_2^2} + \frac{\partial^2 u_\varphi}{\partial x_3^2} = \frac{t}{4\pi} \iint_{S_1} \left(\frac{\partial^2 \varphi}{\partial \alpha_1^2} + \frac{\partial^2 \varphi}{\partial \alpha_2^2} + \frac{\partial^2 \varphi}{\partial \alpha_3^2} \right) d\sigma_1$$
$$= \frac{1}{4\pi t} \iint_{S_t} \left(\frac{\partial^2 \varphi}{\partial \alpha_1^2} + \frac{\partial^2 \varphi}{\partial \alpha_2^2} + \frac{\partial^2 \varphi}{\partial \alpha_3^2} \right) d\sigma_t. \tag{12.8}$$

To calculate $\frac{\partial^2 u_\varphi}{\partial t^2}$, we rewrite equation (12.7) as follows:

$$\frac{\partial u_\varphi}{\partial t} = \frac{u_\varphi}{t} + \frac{1}{4\pi t} \iint_{S_t} \left(\frac{\partial \varphi}{\partial \alpha_1} d\alpha_2\, d\alpha_3 + \frac{\partial \varphi}{\partial \alpha_2} d\alpha_1\, d\alpha_3 + \frac{\partial \varphi}{\partial \alpha_3} d\alpha_1\, d\alpha_2 \right)$$
$$= \frac{u_\varphi}{t} + \frac{1}{4\pi t} \iiint_{V_t} \left(\frac{\partial^2 \varphi}{\partial \alpha_1^2} + \frac{\partial^2 \varphi}{\partial \alpha_2^2} + \frac{\partial^2 \varphi}{\partial \alpha_3^2} \right) d\alpha_1\, d\alpha_2\, d\alpha_3$$
$$= \frac{u_\varphi}{t} + \frac{I(t)}{4\pi t}, \tag{12.9}$$

where

$$I(t)=\iiint\limits_{V_t}\left(\frac{\partial^2\varphi}{\partial\alpha_1^2}+\frac{\partial^2\varphi}{\partial\alpha_2^2}+\frac{\partial^2\varphi}{\partial\alpha_3^2}\right)d\alpha_1\,d\alpha_2\,d\alpha_3,$$

and V_t is the sphere of radius t with centre at the point (x_1, x_2, x_3) on the hyperplane $t=0$.

From formula (12.9), we obtain

$$\begin{aligned}\frac{\partial^2 u_\varphi}{\partial t^2}&=-\frac{u_\varphi}{t^2}+\frac{1}{t}\left[\frac{u_\varphi}{t}+\frac{I(t)}{4\pi t}\right]-\frac{I(t)}{4\pi t^2}+\frac{1}{4\pi t}\frac{\partial I(t)}{\partial t}\\&=\frac{1}{4\pi t}\frac{\partial I(t)}{\partial t}.\end{aligned}\tag{12.10}$$

But it is easy to see that*

$$\frac{\partial I(t)}{\partial t}=\iint\limits_{S_t}\left(\frac{\partial^2\varphi}{\partial\alpha_1^2}+\frac{\partial^2\varphi}{\partial\alpha_2^2}+\frac{\partial^2\varphi}{\partial\alpha_3^2}\right)d\sigma_t\,.\tag{12.11}$$

When we compare equations (12.8), (12.10), and (12.11), we see that the function u defined by Kirchhoff's formula does indeed satisfy the wave equation (12.1).

Remark: If the function $\varphi_1(x_1, x_2, x_3)$ is continuous but not necessarily differentiable and if $\varphi_0(x_1, x_2, x_3)$ and its first derivatives are continuous, the function u defined by equations (12.4) and (12.5) will give only a generalised solution to the Cauchy problem. In this case, a generalised solution to the Cauchy problem for equation (12.1) with initial conditions (12.2) means the limit of a uniformly convergent sequence of solutions $u_{(n)}(t, x_1, x_2, x_3)$ of equation (12.1) with the initial conditions

$$u_{(n)}(0, x_1, x_2, x_3)=\varphi_{0(n)}(x_1, x_2, x_3),$$

$$\frac{\partial}{\partial t}u_{(n)}(0, x_1, x_2, x_3)=\varphi_{1(n)}(x_1, x_2, x_3),$$

* Specifically, if we shift to polar coordinates (ρ, θ, ψ) with centre at the point (x_1, x_2, x_3), we have

$$I(t)=\int_0^t\int_0^\pi\int_0^{2\pi}\Delta\varphi(r, \psi, \theta)\,r^2\sin\theta\,d\psi\,d\theta\,dr,$$

$$\frac{\partial I(t)}{\partial t}=\int_0^\pi\int_0^{2\pi}\Delta\varphi(t, \psi, \theta)\,t^2\sin\theta\,d\psi\,d\theta=\iint\limits_{S_t}\Delta\varphi\,d\sigma_t.$$

if the sequences $\varphi_{0(n)}$, $\frac{\partial\varphi_{0(n)}}{\partial x_i}$, $\varphi_{1(n)}$ converge in G_0 uniformly to φ_0, $\frac{\partial\varphi_0}{\partial x_i}$, φ_1 as $n\to\infty$. It is easy to see that, if $\varphi_1(x_1, x_2, x_3)$ is continuous and φ_0 is continuously differentiable, the generalised solution to the Cauchy problem with initial conditions (12.2) exists and is unique.

2. Let us consider the particular case in which the function φ does not depend on x_3. It is easy to see that the function u given by Kirchhoff's formula will also be independent of x_3 and hence will satisfy the equation

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}. \tag{12.12}$$

In this case, the integral over the sphere S_t can be replaced with a double integral over the intersection K_t of the sphere V_t with the plane $\alpha_3=x_3$. If we project the element $d\sigma_t$ of the surface onto this plane, we obtain

$$d\sigma_t=\frac{t}{\sqrt{t^2-(\alpha_1-x_1)^2-(\alpha_2-x_2)^2}}\,d\alpha_1\,d\alpha_2,$$

and Kirchhoff's formula can be rewritten in the following form

$$u_\varphi(t, x_1, x_2)=\frac{1}{4\pi}\iint_{S_t}\frac{\varphi(\alpha_1, \alpha_2)}{t}\,d\sigma_t$$

$$=\frac{1}{2\pi}\iint_{K_t}\frac{\varphi(\alpha_1, \alpha_2)\,d\alpha_1\,d\alpha_2}{\sqrt{t^2-(\alpha_1-x_1)^2-(\alpha_2-x_2)^2}}.$$

Therefore, the solution to equation (12.12) satisfying the conditions

$$u(0, x_1, x_2)=\varphi_0(x_1, x_2),$$

$$u_t'(0, x_1, x_2)=\varphi_1(x_1, x_2),$$

is given by the formula

$$u(t, x_1, x_2)=\frac{1}{2\pi}\iint_{K_t}\frac{\varphi_1(\alpha_1, \alpha_2)\,d\alpha_1\,d\alpha_2}{\sqrt{t^2-(\alpha_1-x_1)^2-(\alpha_2-x_2)^2}}$$

$$+\frac{1}{2\pi}\frac{\partial}{\partial t}\iint_{K_t}\frac{\varphi_0(\alpha_1, \alpha_2)\,d\alpha_1\,d\alpha_2}{\sqrt{t^2-(\alpha_1-x_1)^2-(\alpha_2-x_2)^2}}. \quad (12.13)$$

This formula is known as Poisson's formula.

3. If the function φ is independent of both x_2 and x_3, the function u given by Kirchhoff's formula will also be independent of x_2 and x_3 and will therefore satisfy

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x_1^2}. \quad (12.14)$$

In this case, Kirchhoff's formula can be rewritten in the form

$$u_\varphi(t, x_1)=\frac{1}{4\pi}\iint_{S_t}\frac{\varphi(\alpha_1)}{t}\,d\sigma_t=\frac{1}{2}\int_{x_1-t}^{x_1+t}\varphi(\alpha_1)\,d\alpha_1.$$

Here, we used the fact* that the area of the portion of the sphere S_t between the intersections of this sphere with the planes $\alpha_1=\text{const}$ and $\alpha_1+d\alpha_1=\text{const}$ is equal to $2\pi t\,d\alpha_1$ and that the function $\varphi(\alpha_1)$ has a constant value on all this portion of the sphere with an accuracy up to magnitudes of order $d\alpha_1$.

Therefore, the solution to equation (12.14) satisfying the conditions

$$u(0, x_1)=\varphi_0(x_1), \qquad u'_t(0, x_1)=\varphi_1(x_1),$$

is given by the formula

$$u(t, x_1)=\frac{1}{2}\int_{x_1-t}^{x_1+t}\varphi_1(\alpha_1)\,d\alpha_1+\frac{1}{2}\frac{\partial}{\partial t}\int_{x_1-t}^{x_1+t}\varphi_0(\alpha_1)\,d\alpha_1$$

$$=\frac{\varphi_0(x_1+t)+\varphi_0(x_1-t)}{2}+\frac{1}{2}\int_{x_1-t}^{x_1+t}\varphi_1(\alpha_1)\,d\alpha_1. \quad (12.15)$$

* The area of a spherical strip of small width $d\alpha$ is approximately equal to $2\pi\rho\,dl$, where ρ is the radius of the middle section of the strip and dl is a generator of the cone passing through the lateral edges of the strip. But $\frac{t}{\rho}=\frac{d}{d\alpha}$ so that $\rho\,dl = t\,d\alpha$ and $d\sigma_t = 2\pi t\,d\alpha$.

This formula is known as d'Alembert's formula.

We recall from the uniqueness theorem proven in Section 11 that there are no solutions to the Cauchy problem besides the ones given for equations (12.1), (12.12), and (12.14) by formulae (12.4), (12.13), and (12.15) respectively. The method that we used to obtain the solution to the Cauchy problem for equations (12.12) and (12.14) from the solution to the Cauchy problem for equation (12.1) is called the method of descent.

We have found the solution to the Cauchy problem for $t>0$. The case of $t<0$ reduces to this case by replacing t with $-t$. This change does not affect equations (12.1), (12.12), and (12.14).

Problem 1. Suppose that $\bar{u}(t, x_1, x_2, x_3; \tau)$ is a solution to equation (12.1) satisfying for $t=\tau$ the conditions

$$\bar{u}(\tau, x_1, x_2, x_3; \tau)=0,$$
$$\frac{\partial \bar{u}}{\partial t}(\tau, x_1, x_2, x_3; \tau)=f(\tau, x_1, x_2, x_3).$$

Show that the solution $u(t, x_1, x_2, x_3)$ of the equation

$$\frac{\partial^2 u}{\partial t^2}-\frac{\partial^2 u}{\partial x_1^2}-\frac{\partial^2 u}{\partial x_2^2}-\frac{\partial^2 u}{\partial x_3^2}=f(t, x_1, x_2, x_3),$$

satisfying at $t=0$ the conditions

$$u(0, x_1, x_2, x_3)=0,$$
$$\frac{\partial u}{\partial t}(0, x_1, x_2, x_3)=0,$$

is given by the formula

$$u(t, x_1, x_2, x_3)=\int_0^t \bar{u}(t, x_1, x_2, x_3; \tau)\, d\tau. \tag{12.16}$$

Problem 2. By using formula (12.5), show that the solution (12.16) is of the form

$$u(t, x_1, x_2, x_3)=\frac{1}{4\pi}\iiint\limits_{r\leqslant t}\frac{f(\alpha_1, \alpha_2, \alpha_3, t-r)}{r}\, d\alpha_1\, d\alpha_2\, d\alpha_3, \tag{12.17}$$

where

$$r=\sqrt{(x_1-\alpha_1)^2+(x_2-\alpha_2)^2+(x_3-\alpha_3)^2}.$$

The integral (12.17) is called a lagging potential.

13. INVESTIGATION OF THE FORMULAE RESOLVING THE CAUCHY PROBLEM

1. The continuous dependence of the solution on the initial conditions. All the formulae giving the solution to the Cauchy problem for the equation

$$\frac{\partial^2 u}{\partial t^2}=\sum_{i=1}^{n}\frac{\partial^2 u}{\partial x_i^2} \tag{13.1}$$

for $n=2$ or 3 that we gave in the preceding section contain integrals of the initial functions multiplied by specific functions and the time derivatives of such integrals. For $n=1$, these formulae contain only integrals of the initial functions and the initial functions themselves.

Therefore, if we replace the initial functions φ_0 and φ_1 in such a way that they and their first derivatives change by a small enough amount, the function $u(t, x_1, \ldots, x_n)$ giving the solution to the Cauchy problem will also change only a slight amount. When $n=1$, for this it is sufficient that the functions φ_0 and φ_1 change only slightly. In this case, it is of course assumed that the values of t in question are bounded if the region in which the initial functions are given is infinite.

Thus, it is established that the Cauchy problem for equations (12.1), (12.12), and (12.14) is correctly posed.

It is possible to derive formulae* giving the solution to the Cauchy problem for equation (13.1) for arbitrary n that are analogous to formulae (12.4), (12.13), and (12.15) and to show that the Cauchy problem is correctly stated in this case also if the initial conditions are given for $t=0$. The numbers L_1 and L_2 that appear in the definition of correctness (see Section 8) are equal respectively to $[n/2]+2$ and $[n/2]$. Here, $[x]$ denotes the greatest integer not exceeding x.

It follows from formulae (12.4) and (12.13) that, for small values of t, the quantities

$$|u(t, x_1, x_2, x_3)| \text{ and } |u(t, x_1, x_2)|$$

* These formulae may, for example, be derived by the method expounded in *Course in Higher Mathematics* by SMIRNOV, V.I., **II**, Section 173, Fizmatgiz (1958).

respectively can be quite great despite the smallness of φ_0 and φ_1 if the derivatives of the function φ_0 are large. Wave 'splashes' may be formed.

2. The diffusion of waves. Formulae (12.4) and (12.5) show that the value of the solution to the Cauchy problem for the wave equation (13.1) with $n=3$ at the point $(t, x_1, \ldots, x_n)$ depends on the initial conditions only on the boundary of the base of the characteristic cone whose vertex is at the point (t, x_1, x_2, x_3). On the other hand, if n is 1 or 2, then $u(t, x_1, \ldots, x_n)$ depends on the initial data throughout the entire base of this cone, as is shown by formulae (12.13) and (12.15).

Let us suppose that the initial values of u and u'_t are nonzero at $t=0$ only within a small region G_ε around some point $(0, x_1^0, \ldots, x_n^0)$. Let us investigate the values of u at the points $(t, x_1, \ldots, x_n)$ for fixed values of $x_1, \ldots, x_n$ and for increasing values of t (beginning with 0). For $n=3$, the value of $u(t, x_1, \ldots, x_n)$ can be nonzero only on a small segment of a straight line parallel to the t-axis in the space $(t, x_1, \ldots, x_n)$, specifically, on that segment containing the vertices of the characteristic cones of equation (12.1) the boundaries of whose bases intersect the region G_ε. On the other hand, if n is 1 or 2 and if the point $(0, x_1)$ (resp. $(0, x_1, x_2)$) does not belong to G_ε, then $u(t, x_1)$ (resp. $u(t, x_1, x_2)$) will be equal to zero for sufficiently small values of t and it will generally assume nonzero values beginning with those values of t at which the segment $|x_1 - \alpha_1| \leq t$ (resp. $(\alpha_1 - x_1)^2 + (\alpha_2 - x_2)^2 \leq t^2$) intersects the region G_ε.

Consequently, a disturbance that is set up at the initial moment in some small neighbourhood of the point $(x_1^0, \ldots, x_n^0)$ is, for $n=3$ and $t>0$, reflected in the values of the function only at those points in the space $(x_1, \ldots, x_n)$ that lie close to the sphere of radius t with centre at the point $(x_1^0, \ldots, x_n^0)$. Thus, a disturbance set up at the initial instant at the point (x_1^0, x_2^0, x_3^0) produces a spherical wave having centre at that point and possessing a leading and a trailing edge. On the other hand, if n is 1 or 2, a disturbance set up at the initial instant in a neighbourhood of the point $(x_1^0, \ldots, x_n^0)$ generally has an effect at all points lying within the sphere of radius t with centre at $(x_1^0, \ldots, x_n^0)$. A wave possessing a sharp leading edge and a dissolving trail-

ing edge is set up. In such a case, we say that diffusion of the wave (dissolution of the trailing edge) takes place. For $n=3$, there is no diffusion. It can be shown that there is no diffusion of waves for solutions of equation (13.1) if n is any odd number greater than 1.

Disturbances set up in a small region G_ε of a three-dimensional elastic body or gas cause waves that leave no trace behind them if we assume that their vibrations obey equation (12.1). In the case of a gas, for example, $u(t, x_1, x_2, x_3)$ denotes the deviation in the pressure of the gas at the point (x_1, x_2, x_3) at the instant t from the normal pressure. On the other hand, disturbances in a two-dimensional continuum (for example, of a stretched membrane or a liquid surface) that are set up in a small region G_ε produce waves that theoretically always leave a trace behind if we assume that these vibrations obey equation (12.12). Actually, these vibrations are damped very rapidly as a result of friction, which is not taken into consideration in the derivation of equation (12.12). Similarly, a trace generally remains after a wave passes through a point in a one-dimensional continuum (see subsection 3 of the present section).

3. Examination of d'Alembert's formula. Let us consider two particular cases that show clearly the behaviour of the solution of equation (12.14) in the general case.

First, let us consider the case in which $\varphi_1(x)\equiv 0$ and the graph of $\varphi_0(x)$ has the form shown at the top of Fig. 3 (the heavy continuous line). For brevity, we shall write x instead of x_1. Then, d'Alembert's formula takes the form

$$u(t, x)=\frac{\varphi_0(x+t)+\varphi_0(x-t)}{2}.$$

To get a graph of $u(t, x)$, which is considered as a function of x at some fixed positive value of t, it is convenient to proceed as follows: First, we draw two identical coincident graphs each of which is obtained from the graph $\varphi_0(x)$ by halving the ordinate at every value of x (the dashed line at the top of Fig. 3). Then, we displace one of these graphs without changing its shape (just as if it were a rigid body) by a distance t to the right along the positive half of the x-axis, and we move the other one by an amount t to the left. Then, we must construct a new graph whose ordinate at every value

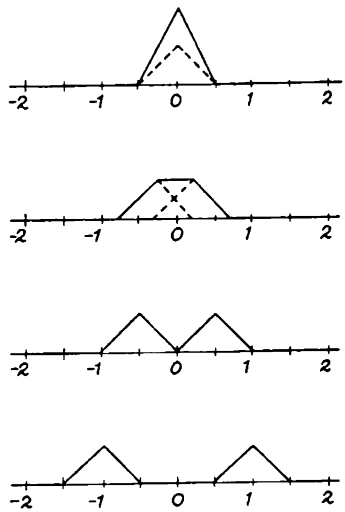


Fig. 3

of x is equal to the sum of the ordinates of the same x of the two moving graphs. The drawing shows the graphs constructed in this manner for

$$u(0, x),\ u\left(\frac{1}{4}, x\right),$$
$$u\left(\frac{1}{2}, x\right),\ u(1, x)$$

(Throughout the drawing, the dashed lines represent the auxiliary graphs and the continuous heavy line represent the graphs of $u(t, x)$ for some fixed value of t.)

Let us now consider the case in which $\varphi_0(x) \equiv 0$ and

$$\varphi_1(x) = \begin{cases} 1 & \text{for } |x| \leqslant \frac{1}{2}, \\ 0 & \text{for } |x| > \frac{1}{2}. \end{cases}$$

Then, d'Alembert's formula takes the form

$$u(t, x) = \frac{1}{2} \int_{x-t}^{x+t} \varphi_1(\alpha)\, d\alpha.$$

For every fixed value of x, the value of $u(t, x)$ will be 0 until the interval $(x-t, x+t)$ contains the interval (-1/2, 1/2), where $\varphi_1(x) \neq 0$. The value of $u(t, x)$ will vary during

an interval of time when the lengthened interval $(x-t, x+t)$ will contain an ever larger portion of the interval (-1/2, 1/2). After the interval $(x-t, x+t)$ contains the entire interval (-1/2, 1/2), the value of $u(t, x)$ will keep the constant value

$$\frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \varphi_1(\alpha)\,d\alpha.$$

To obtain a graph representing the shape of the string at different values of t, we proceed as follows:

We denote by $\Phi(z)$ some primitive function of $\varphi_1(z)$. Then,

$$u(t, x) = \frac{1}{2}[\Phi(x+t) - \Phi(x-t)].$$

To obtain the graph $u(t, x)$, we draw the graphs of the functions $\Phi(x)/2$ and $-\Phi(x)/2$. Then, we displace each of these graphs (as before, without changing their shapes) by an amount t along the x-axis, the first to the left and the second to the right. When we add the ordinates of the displaced graphs, we obtain the graph of the function $u(t, x)$.

Figure 4 shows the shape of the string at the instants $t=0$, 1/4, 1/2, and 1.

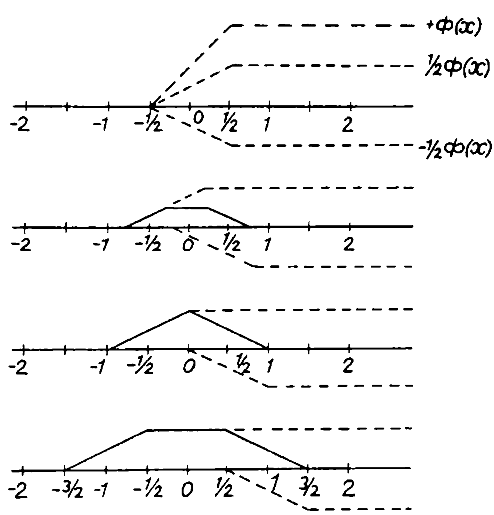


Fig. 4

The phenomenon of diffusion is expressed here in the fact that the point x, as it moves away from the equilibrium position, does not again return to it.

The functions $\varphi_0(x)$ and $\varphi_1(x)$ that we have been examining in these examples either possess discontinuities (as is the case with $\varphi_1(x)$) or have derivatives possessing discontinuities (as is the case with $\varphi_0(x)$). Therefore, the generalised solutions of equation (12.14) do not correspond to them. To obtain an ordinary twice continuously differentiable solution of this equation, we need only make a slight change in the graphs of the functions $\varphi_0(x)$ and $\varphi_1(x)$ so as to obtain the graphs of functions with continuous second derivatives. For the function φ_0, this can be done in such a way that the ordinate $\varphi_0(x)$ will change everywhere only slightly. Then, the corresponding solution of equation (12.14) will also change everywhere only slightly. The replacement of $\varphi_1(x)$ with a continuous smooth function can be done in such a way that $\Phi(x)$ would change by an arbitrarily small amount. Then $u(t, x)$ would change only slightly everywhere.

14. THE LORENTZ TRANSFORMATIONS

1. In Section 1, we pointed out that the expression

$$\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}$$

is, up to a constant factor, the only linear combination of second derivatives that remains invariant in form when the space is rotated, that is, when an arbitrary transformation of the coordinates x_1, x_2, x_3 is made. The wave equation

$$\frac{\partial^2 u}{\partial t^2}-\frac{\partial^2 u}{\partial x_1^2}-\frac{\partial^2 u}{\partial x_2^2}-\frac{\partial^2 u}{\partial x_3^2}=0 \tag{14.1}$$

is also very closely associated with a certain class of linear transformations of the variables (t, x_1, x_2, x_3) with real constant coefficients, which do not change the form of this equation. Let us look at them in greater detail.

Any linear homogeneous transformation of these variables with real coefficients of the form

$$y_i=\sum_{j=0}^{3} a_{ij}x_j \quad (i=0,\ 1,\ 2,\ 3), \tag{14.2}$$

that leaves the quadratic form

$$x_0^2 - x_1^2 - x_2^2 - x_3^2 \tag{14.3}$$

an invariant, that yields new variables of the form

$$y_0^2 - y_1^2 - y_2^2 - y_3^2$$

is known as a Lorentz transformation.

It is easy to show that the set of all Lorentz transformations constitutes a group, the group operation of which is the combination of the successive transformations. In particular, it is easy to see that successive application of two Lorentz transformations always yields a Lorentz transformation.

Let us write the formula for some particular class of Lorentz transformations. Let us consider a transformation that leaves invariant two out of the last three (space) coordinates. Such a transformation is of the form

$$\left.\begin{aligned} y_0 &= ax_0 + bx_1, \\ y_1 &= cx_0 + dx_1, \\ y_2 &= \qquad\qquad x_2, \\ y_3 &= \qquad\qquad\quad x_3. \end{aligned}\right\} \tag{14.4}$$

With this transformation, the identity

$$y_0^2 - y_1^2 \equiv x_0^2 - x_1^2.$$

must be satisfied. When we substitute y_0 and y_1 from formulae (14.4), we have

$$(ax_0 + bx_1)^2 - (cx_0 + dx_1)^2 \equiv x_0^2 - x_1^2.$$

Therefore,

$$\left.\begin{aligned} a^2 - c^2 &= 1, \\ b^2 - d^2 &= -1, \\ ab - cd &= 0. \end{aligned}\right\} \tag{14.5}$$

These equations will be satisfied if we set

$$a = d = \frac{1}{\sqrt{1-\beta^2}}, \quad b = c = \frac{\beta}{\sqrt{1-\beta^2}},$$

where $|\beta| < 1$.

We thus obtain formulae for a certain class of Lorentz transformations:

$$\left.\begin{aligned} y_0 &= \frac{x_0+\beta x_1}{\sqrt{1-\beta^2}}, \\ y_1 &= \frac{\beta x_0 + x_1}{\sqrt{1-\beta^2}}, \\ y_2 &= \qquad x_2, \\ y_3 &= \qquad\quad x_3. \end{aligned}\right\} \tag{14.6}$$

Formulae (14.6) are quite significant since we shall now show that every Lorentz transformation is a combination of (1) an orthogonal transformation of the variables x_1, x_2 and x_3 that leaves x_0 invariant, (2) a transformation of the form (14.6), and (3) a change of sign of some of the variables (reflection).

Suppose that a Lorentz transformation is given by the formulae

$$\left.\begin{aligned} y_0 &= a_{00}x_0 + a_{01}x_1 + a_{02}x_2 + a_{03}x_3, \\ y_1 &= a_{10}x_0 + a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ y_2 &= a_{20}x_0 + a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ y_3 &= a_{30}x_0 + a_{31}x_1 + a_{32}x_2 + a_{33}x_3. \end{aligned}\right\} \tag{14.7}$$

If at least one of the numbers a_{01}, a_{02}, and a_{03} is nonzero, let us perform an orthogonal transformation of x_1, x_2, x_3 into x_1', x_2', x_3' such that

$$a_{01}x_1 + a_{02}x_2 + a_{03}x_3 = ax_1'.$$

If in addition x_0' is set equal to x_0, then, as is easily seen, this transformation from x_0, x_1, x_2, x_3 to x_0', x_1', x_2', x_3' is a Lorentz transformation. If we substitute the variables x_0', x_1', x_2', and x_3' into the right side of the system (14.7), we obtain

$$\left.\begin{aligned} y_0 &= a_{00}x_0' + ax_1', \\ y_1 &= a_{10}x_0' + b_{11}x_1' + b_{12}x_2' + b_{13}x_3', \\ y_2 &= a_{20}x_0' + b_{21}x_1' + b_{22}x_2' + b_{23}x_3', \\ y_3 &= a_{30}x_0' + b_{31}x_1' + b_{32}x_2' + b_{33}x_3'. \end{aligned}\right\} \tag{14.8}$$

Let us show that $a^2 < a_{00}^2$. First, since (14.8) is a Lorentz transformation, we have

$$y_0^2 - y_1^2 - y_2^2 - y_3^2 = x_0'^2 - x_1'^2 - x_2'^2 - x_3'^2,$$

so that

$$y_1^2+y_2^2+y_3^2=x_1'^2+x_2'^2+x_3'^2-x_0'^2+y_0^2. \tag{14.9}$$

Let us set $y_0=0$. Then,

$$x_0'=-\frac{a}{a_{00}}x_1'$$

and the identity (14.9) becomes an identity in the three variables:

$$y_1^2+y_2^2+y_3^2=\left(1-\frac{a^2}{a_{00}^2}\right)x_1'^2+x_2'^2+x_3'^2.$$

The right side is positive for arbitrary values of x_1', x_2', and x_3' if $x_1'^2+x_2'^2+x_3'^2>0$ since the fact that $y_0=y_1=y_2=y_3=0$ implies that $x_0=x_1=x_2=x_3=0$. Therefore,

$$1-\frac{a^2}{a_{00}^2}>0,$$

or $a^2<a_{00}^2$.

Let us set $\frac{a}{a_{00}}=\beta$ and let us perform a Lorentz transformation of the form (14.6):

$$\left.\begin{aligned} x_0''&=\frac{x_0'+\beta x_1'}{\sqrt{1-\beta^2}},\\ x_1''&=\frac{\beta x_0'+x_1'}{\sqrt{1-\beta^2}},\\ x_2''&=x_2',\\ x_3''&=x_3'. \end{aligned}\right\} \tag{14.10}$$

Obviously, y_0, y_1, y_2, and y_3 will be connected with x_0'', x_1'', x_2'', and x_3'' by a Lorentz transformation of the form

$$\left.\begin{aligned} y_0&=cx_0'',\\ y_1&=c_{10}x_0''+c_{11}x_1''+c_{12}x_2''+c_{13}x_3'',\\ y_2&=c_{20}x_0''+c_{21}x_1''+c_{22}x_2''+c_{23}x_3'',\\ y_3&=c_{30}x_0''+c_{31}x_1''+c_{32}x_2''+c_{33}x_3'', \end{aligned}\right\} \tag{14.11}$$

where
$$c=\pm\sqrt{a_{00}^2-a^2}$$

(as can easily be verified).

If $a_{01}=a_{02}=a_{03}=0$, the system (14.7) already has the form (14.11).

Let us find the values of the coefficients c, c_{10}, c_{20}, and c_{30}. If we set $x''_0=1$ and $x''_1=x''_2=x''_3=0$, we obtain,

$$y_0=c,\ y_1=c_{10},\ y_2=c_{20},\ y_3=c_{30}.$$

Therefore,

$$1=c^2-c_{10}^2-c_{20}^2-c_{30}^2 \text{ and } c^2\geqslant 1.$$

If we set $y_0=1$ and $y_1=y_2=y_3=0$, we find that $x''_0=\frac{1}{c}$ and that x''_1, x''_2, and x''_3 have certain definite values $\breve{x}''_1, \breve{x}''_2$, and $\breve{x}''_3$. Therefore,

$$1=\frac{1}{c^2}-\breve{x}''^2_1-\breve{x}''^2_2-\breve{x}''^2_3 \text{ and } \frac{1}{c^2}\geqslant 1,$$

or $c^2\leqslant 1$.

Consequently, $c^2=1$, and, returning to the equation

$$1=c^2-c_{10}^2-c_{20}^2-c_{30}^2,$$

we see that $c_{10}=c_{20}=c_{30}=0$. Consequently, the transformation (14.11) does indeed have the form

$$\left.\begin{aligned} y_0&=\pm x''_0,\\ y_1&=\qquad c_{11}x''_1+c_{12}x''_2+c_{13}x''_3,\\ y_2&=\qquad c_{21}x''_1+c_{22}x''_2+c_{23}x''_3,\\ y_3&=\qquad c_{31}x''_1+c_{32}x''_2+c_{33}x''_3. \end{aligned}\right\} \qquad (14.12)$$

By changing the sign of the coordinate x''_0 if necessary, we obtain a Lorentz transformation that is a simple orthogonal transformation of the variables x''_1, x''_2, and x''_3 into y_1, y_2, and y_3.

Thus, we see that the most general Lorentz transformation (14.7) that maps the variables x_i into y_i is the result of successive transformations: an orthogonal transformation mapping x_i into x'_i, a Lorentz transformation of the particular form (14.6) mapping x'_i into x''_i, possibly a change of sign of x''_0, and, finally, an orthogonal transformation of x''_i into y_i (for $i=1, 2, 3$).

If we transpose the matrix of each of these intermediary transformations, we again obtain the matrix of a transformation of the same type. It follows from this that the trans-

pose of the matrix of a Lorentz transformation is also the matrix of a Lorentz transformation. Furthermore, it follows from the definition of a Lorentz transformation that the inverse of a Lorentz transformation is also a Lorentz transformation.

2. Let us now show the basic fact which will make clear the close connection between Lorentz transformations and the wave equation.

Theorem. Every nonsingular linear transformation of the variables t, x_1, x_2, x_3 with real constant coefficients, which does not change the form of equation (14.1), is a combination of a Lorentz transformation, a translation of the coordinate origin in the space (t, x_1, x_2, x_3), and a similarity transformation in that space.

For simplicity in notation, we set $t = x_0$.

The assertion that some transformation 'does not change the form of the equation' is to be understood as follows: An arbitrary function $u(x_0, x_1, x_2, x_3)$ (with continuous second derivatives) that satisfies the equation

$$\frac{\partial^2 u}{\partial x_0^2} - \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} - \frac{\partial^2 u}{\partial x_3^2} = 0,$$

will, after a transformation of the x_i into y_i, be transformed into a function $u(y_0, y_1, y_2, y_3)$ satisfying the equation

$$\frac{\partial^2 u}{\partial y_0^2} - \frac{\partial^2 u}{\partial y_1^2} - \frac{\partial^2 u}{\partial y_2^2} - \frac{\partial^2 u}{\partial y_3^2} = 0. \tag{14.13}$$

From this, it follows that under an arbitrary transformation of this sort for an arbitrary function $u(x_0, x_1, x_2, x_3)$ the equation

$$\frac{\partial^2 u}{\partial y_0^2} - \frac{\partial^2 u}{\partial y_1^2} - \frac{\partial^2 u}{\partial y_2^2} - \frac{\partial^2 u}{\partial y_3^2} = k\left(\frac{\partial^2 u}{\partial x_0^2} - \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} - \frac{\partial^2 u}{\partial x_3^2}\right), \tag{14.14}$$

where $k \neq 0$ is a constant, will be valid. To see this, if we make the most general assumption

$$\frac{\partial^2 u}{\partial y_0^2} - \frac{\partial^2 u}{\partial y_1^2} - \frac{\partial^2 u}{\partial y_2^2} - \frac{\partial^2 u}{\partial y_3^2} \equiv \sum_{i,j=0}^{3} A_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} \tag{14.15}$$

and suppose that

$$\sum_{i,j=0}^{3} A_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} \neq k\left(\frac{\partial^2 u}{\partial x_0^2}-\frac{\partial^2 u}{\partial x_1^2}-\frac{\partial^2 u}{\partial x_2^2}-\frac{\partial^2 u}{\partial x_3^2}\right),$$

we arrive at a contradiction with the fact that every solution to equation (14.1) is transformed into a solution of the same equation under a transformation of the variable. In this case, we may choose a system of numbers $u^0_{ik}=u^0_{ki}$, that satisfies the two linear equations

$$\sum_{i,j=0}^{3} A_{ij}u^0_{ij}=1, \tag{14.16$_1$}$$

$$u^0_{00}-u^0_{11}-u^0_{22}-u^0_{33}=0. \tag{14.16$_2$}$$

For the function

$$u(x_0, x_1, x_2, x_3)=\frac{1}{2}\sum_{i,j=0}^{3} u^0_{ij}x_ix_j$$

we have the equations

$$\frac{\partial^2 u}{\partial x_i \partial x_j}=u^0_{ij}.$$

On the basis of $(14.16)_2$, this function satisfies equation (14.1) and, after a change of variables, it will satisfy equation (14.13). This follows from $(14.16)_1$ and (14.15). Consequently, equation (14.14) is valid.

Let us perform the similarity transformation

$$x'_i=x_i\frac{1}{\sqrt{|k|}}\ (i=0,\ 1,\ 2,\ 3);$$

Then,

$$k\left(\frac{\partial^2 u}{\partial x_0^2}-\frac{\partial^2 u}{\partial x_1^2}-\frac{\partial^2 u}{\partial x_2^2}-\frac{\partial^2 u}{\partial x_3^2}\right)=\pm\left(\frac{\partial^2 u}{\partial x_0'^2}-\frac{\partial^2 u}{\partial x_1'^2}-\frac{\partial^2 u}{\partial x_2'^2}-\frac{\partial^2 u}{\partial x_3'^2}\right).$$

Consequently, it will be sufficient for us to show that the transformation

$$y_i=\sum_{j=0}^{3} a_{ij}x'_j \qquad (i=0,\ \ldots,\ 3), \tag{14.17}$$

which does not change the absolute value of the differential expression

$$\frac{\partial^2 u}{\partial x_0^{2}} - \frac{\partial^2 u}{\partial x_1'^{2}} - \frac{\partial^2 u}{\partial x_2'^{2}} - \frac{\partial^2 u}{\partial x_3'^{2}}, \tag{14.18}$$

is a Lorentz transformation; that is, it does not change the form of the quadratic form

$$x_0'^2 - x_1'^2 - x_2'^2 - x_3'^2. \tag{14.19}$$

But this follows from the fact, that, as was shown in Section 5, under a linear transformation of the independent variables of the form (14.17), the expression

$$\sum_{i,j=0}^{3} c_{ij} \frac{\partial^2 u}{\partial x_i' \partial x_j'}$$

is transformed in the same way as the quadratic form composed of these variables

$$\sum_{i,j=0}^{3} c_{ij} x_i' x_j',$$

is transformed if we perform the transformation

$$x_j' = \sum_{i=0}^{3} a_{ij} y_i \qquad (j = 0, \ldots, 3). \tag{14.20}$$

on them. Since the transformation (14.17) does not change the form of the expression (14.18) except as regards sign, the transformation (14.20) does not change the form of the quadratic form (14.19) except as regards sign. But, on the basis of the law of inertia, even the sign of (14.19) cannot change under any linear transformation with real coefficients. Therefore, the transformation (14.20) and its inverse are Lorentz transformations. In accordance with what was shown in subsection 1, the original transformation (14.17) must therefore be a Lorentz transformation since its matrix is the transpose of the matrix of the Lorentz transformation (14.20).

Thus, we have shown that every homogeneous linear transformation that does not change the form of equation (14.1) is a combination of a similarity transformation and a Lorentz transformation. Since a translation of the coordinate origin

obviously does not change the form of this equation, the theorem is proved.

3. By an orthogonal transformation of the variables x_1, x_2, x_3 we can translate an arbitrary hyperplane in the space (t, x_1, x_2, x_3) inclined at an angle greater than 45° from the t-axis (and only such an angle) that passes through the coordinate origin into the hyperplane

$$t = \beta x_1, \text{ where } |\beta| < 1^{\dagger}$$

but the Lorentz transformation (14.6) makes it possible to translate this hyperplane into the coordinate hyperplane $t^* = 0$. Thus, by a linear transformation of the independent variables that does not change the form of equation (14.1), we can always translate an arbitrary hyperplane in the space (t, x_1, x_2, x_3) that is inclined at an angle greater than 45° to the t-axis into the hyperplane $t = 0$. This shows that it is possible to solve the Cauchy problem for equation (14.1) by giving the initial conditions not only on the hyperplane $t = 0$ but on an arbitrary hyperplane Π that forms an angle greater than 45° with the t-axis or, what amounts to the same thing, on a hyperplane Π that intersects each of the characteristic cones of equation (14.1) only along its field or only at its vertex. Specifically, by defining a function u and its derivative in some direction out of the plane Π throughout some region G_0 contained in Π, we give the first derivatives of u in an arbitrary direction in the space (t, x_1, x_2, x_3) throughout the plane G_0 since knowledge of the function u in the

† Suppose that the equation of such a hyperplane is given in the form $At + Bx_1 + Cx_2 + Dx_3 = 0$, where $B^2 + C^2 + D^2 = 1$. Then, the cosine of the angle α_0 between the normal to the hyperplane and the t-axis is equal to $A/(A^2+1)^{1/2}$, and the tangent of this angle is $1/A$. If the normal to the hyperplane makes an angle less than 45° with the t-axis, the transformation

$$Bx_1 + Cx_2 + Dx_3 = x_1'$$

for values of x_2' and x_3' that are suitably chosen (from the conditions of orthogonality of the transformation) transforms the given hyperplane into a hyperplane of the form

$$At + x_1' = 0 \quad \text{or} \quad t = -\frac{1}{A}x_1', \text{where} \left|\frac{1}{A}\right| < 1.$$

region G_0 tells us its first derivatives in every direction in G_0. By transforming the hyperplane Π into the hyperplane $t^* = 0$, we reduce the solution of the Cauchy problem under the initial conditions on Π to the Cauchy problem that we examined in Section 12.

On the other hand, it is easy to show that the Cauchy problem for equation (14.1) will be incorrectly stated if the initial conditions are given on a hyperplane Π that makes an angle not exceeding 45° with the t-axis in the space $t, x_1, \ldots, x_n$. To see this, note that if the hyperplane Π makes a 45° angle with the t-axis, it will have a characteristic direction, and therefore we cannot give Cauchy conditions arbitrarily on it no matter what smoothness requirements we impose upon them.

Let us now consider the case in which Π makes an angle less than 45° with the t-axis. By an orthogonal transformation of coordinates in the space (x_1, x_2, x_3) and by a parallel displacement of them, we can always arrange for the hyperplane Π to have the equation

$$\beta t' + x_1' = 0, \quad \text{where} \quad |\beta| < 1.$$

Here, as we have noted, the form of equation (14.1) is not changed. If we now perform a Lorentz transformation, we can arrange for the hyperplane Π to have the equation

$$x_1^* = 0.$$

This does not change equation (14.1).

Let us give the following Cauchy conditions on the hyperplane $x_1^* = 0$:

$$\left.\begin{aligned} u(t^*, 0, x_2^*, x_3^*) &= \varphi_0(x_2^*), \\ u'_{x_1^*}(t^*, 0, x_2^*, x_3^*) &= \varphi_1(x_2^*). \end{aligned}\right\} \tag{14.21}$$

If we find the solution $u(x_1^*, x_2^*)$ of the equation

$$\frac{\partial^2 u}{\partial x_1^{*2}} + \frac{\partial^2 u}{\partial x_2^{*2}} = 0,$$

satisfying the conditions

$$\left.\begin{aligned} u(0, x_2^*) &= \varphi_0(x_2^*), \\ u'_{x_1^*}(0, x_2^*) &= \varphi_1(x_2^*), \end{aligned}\right\} \tag{14.22}$$

the function $u(x_1^*, x_2^*)$ will satisfy equation (14.1) and the

conditions (14.21). If we take conditions $(8.2)_1$ and $(8.2)_2$, which we used in Hadamard's example, for the initial conditions (14.22), we can easily obtain an incorrectly stated Cauchy problem for equation (14.1) with initial conditions on the hyperplane $x_1 = 0$.

15. MATHEMATICAL BASES OF THE SPECIAL THEORY OF RELATIVITY

The special principle of relativity consists in the fact that all the laws of nature have the same form in all inertial systems (i.e., systems consisting of three space coordinates denoting position and one time coordinate)*. More precisely, all the laws of nature can be written by the same equations in each of these systems. In particular, the velocity of light is the same in each of these systems and hence does not depend on the direction of propagation of the light. For simplicity in writing, we shall assume that this velocity is equal to 1.

Such a system of coordinates is called an inertial system if every body moves in a straight line with uniform velocity in such a system when there are no external forces. It follows from this definition that a space-time system that is moving with a uniform velocity in a straight line with respect to any inertial space-time system is also an inertial system and, conversely, any two inertial systems move with a uniform velocity in a straight line with respect to each other.

Our purpose is to find the relationship between space-time coordinates for two inertial systems A' and A'' one of which A'' is moving uniformly in a straight line with velocity β (where $|\beta| < 1$) with respect to the other inertial system A'.

From the assumption that space and time are homogeneous and isotropic, we shall assume that the connection that we are looking for is linear and that the coefficients depend only on β. We shall denote the space-time coordinates for A' by (t', x_1', x_2', x_3') and for A'' by $(t'', x_1'', x_2'', x_3'')$. For simplicity in writing, we shall sometimes write x_0' instead of t' and x_0'' instead of t''.

Thus, suppose that

* The calculation system is made up of space coordinates which determine the place and time.

$$x_i'' = \sum_{j=0}^{3} a_{ij}(\beta)\, x_j' + \alpha_i \qquad (i = 0, 1, 2, 3). \tag{15.1}$$

The search for the relationship between the coordinates (t', x_1', x_2', x_3') and $(t'', x_1'', x_2'', x_3'')$ will be based only on the constancy of the velocity of light for the systems A' and A''.

A rectilinear propagation of a plane light wave in the space (x_1', x_2', x_3') can be described by a nonconstant function

$$f(a_0 t' + a_1 x_1' + a_2 x_2' + a_3 x_3'), \tag{15.2}$$

the level surfaces of which are displaced perpendicular to the plane

$$a_1 x_1' + a_2 x_2' + a_3 x_3' = \text{const}$$

with changing time t' with a velocity

$$-\frac{a_0}{\sqrt{a_1^2 + a_2^2 + a_3^2}},$$

which by assumption is equal to 1. Here, a_0, a_1, a_2, and a_3 are constant. From this, it follows that

$$a_0^2 = a_1^2 + a_2^2 + a_3^2. \tag{15.3}$$

Since the velocity of light for the system A'' in the coordinates t'', x_1'', x_2'', x_3'' must also be equal to 1, when we shift from the coordinates x_i' to the coordinates x_i'' we see that the expression

$$a_0 t' + a_1 x_1' + a_2 x_2' + a_3 x_3'$$

becomes

$$a_0' t'' + a_1' x_1'' + a_2' x_2'' + a_3' x_3'' + b$$

and that

$$a_0'^2 = a_1'^2 + a_2'^2 + a_3'^2. \tag{15.4}$$

Let us show that the coordinates t'', x_1'', x_2'', x_3'' can be obtained from t', x_1', x_2', x_3' by a Lorentz transformation and a translation of the coordinate origin. By a translation of the origin, we may replace the coordinates t'', x_1'', x_2'', x_3'' with coordinates t, x_1, x_2, x_3 that are related to t', x_1', x_2', x_3' by the homogeneous linear equations

$$x_i = \sum_{j=0}^{3} a_{ij}(\beta) x_j' \qquad (i=0,\ 1,\ 2,\ 3). \tag{15.5}$$

Suppose now that the function

$$f(a_0 x_0' + a_1 x_1' + a_2 x_2' + a_3 x_3')$$

is transformed into the function

$$f(a_0' x_0 + a_1' x_1 + a_2' x_2 + a_3' x_3)$$

If the numbers a_0, a_1, a_2, and a_3 satisfy the equation $a_0^2 - a_1^2 - a_2^2 - a_3^2 = 0$, then a_0', a_1', a_2', a_3' satisfy the analogous equation $a_0'^2 - a_1'^2 - a_2'^2 - a_3'^2 = 0$. Here (a_0, a_1, a_2, a_3) represents an arbitrary system of numbers that satisfies equation (15.3) and a_0', a_1', a_2', a_3' represents the corresponding system of numbers after the transformation (15.5). Let us show that this implies that (15.5) gives the Lorentz transformation for the coefficients a_i; that is,

$$a_0^2 - a_1^2 - a_2^2 - a_3^2 = a_0'^2 - a_1'^2 - a_2'^2 - a_3'^2.$$

For a transformation of the variables a_i when the substitution (15.5) is made, we have in general the formula

$$a_0^2 - a_1^2 - a_2^2 - a_3^2 \equiv \sum_{i,j=0}^{3} k_{ij}(\beta) a_i' a_j'. \tag{15.6}$$

Let us first show that

$$\sum_{i,j=0}^{3} k_{ij}(\beta) a_i' a_j' \equiv k(\beta)(a_0'^2 - a_1'^2 - a_2'^2 - a_3'^2). \tag{15.7}$$

This is true because the equation

$$\sum_{i,j=0}^{3} k_{ij}(\beta) a_i' a_j' = 0 \tag{15.8}$$

implies that

$$a_0'^2 - a_1'^2 - a_2'^2 - a_3'^2 = 0, \tag{15.9}$$

and conversely; that is, the surfaces in the four-dimensional space (a_0', a_1', a_2', a_3') defined by equations (15.8) and (15.9) must coincide; it is then easy to show that formula (15.7)

holds. Consequently,

$$a_0^2 - a_1^2 - a_2^2 - a_3^2 = k(\beta)(a_0'^2 - a_1'^2 - a_2'^2 - a_3'^2).$$

If we consider the motion of the first system relative to the second, which must have a velocity of $-\beta$, we obtain analogously

$$a_0'^2 - a_1'^2 - a_2'^2 - a_3'^2 = k(-\beta)(a_0^2 - a_1^2 - a_2^2 - a_3^2),$$

so that

$$k(\beta)\cdot k(-\beta) = 1.$$

On the other hand, because of the symmetry of the two systems, $k(\beta) = k(-\beta)$. Consequently $k(\beta) = \pm 1$.

Since the transformation (15.5) on the variables x_i' also subjects the variables x_i to a linear transformation, the number of plus signs and minus signs in the quadratic form of the a_i cannot change. Therefore, $k(\beta) = 1$ and the form $a_0^2 - a_1^2 - a_2^2 - a_3^2$ cannot change under the transformation (15.5). Consequently, this transformation of the variables a_i is a Lorentz transformation. The linear transformation to which the variables a_i are subjected when the transformation (15.5) is performed on the x_i is given by the inverse transpose of the matrix (15.5). But then, even the transformation (15.5) is also a Lorentz transformation (see end of subsection 1, Section 14), which was to be proven.

16. SURVEY OF THE BASIC PRINCIPLES OF THE THEORY OF THE CAUCHY PROBLEM. SOME INVESTIGATIONS OF GENERAL HYPERBOLIC EQUATIONS

Up to the present, we spoke of the Cauchy problem for the wave equation (12.1). This section gives a survey of the basic principles in the theory of the Cauchy problem for general hyperbolic equations. We shall omit the proofs and shall consider primarily linear second-order equations.

1. We shall say that the linear equation

$$\frac{\partial^2 u}{\partial t^2} = \sum_{i,j=1}^{n} A_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} A_{0i}\frac{\partial^2 u}{\partial t\, \partial x_i} + \sum_{i=1}^{n} B_i \frac{\partial u}{\partial x_i} + B_0 \frac{\partial u}{\partial t} + Cu + D, \tag{16.1}$$

where the coefficients A_{ij}, A_{0i}, B_i, B_0, C and D are functions of t, $x_1, \ldots, x_n$, is t-hyperbolic in some region G of the space $(t, x_1, \ldots, x_n)$ if the following condition is satisfied: Every straight line passing through the coordinate origin in the real space $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ intersects the surface

$$1 = \sum_{i,j=1}^{n} A_{ij}(t, x_1, \ldots, x_n)\alpha_i\alpha_j + \sum_{i=1}^{n} A_{0i}(t, x_1, \ldots, x_n)\alpha_i \tag{16.2}$$

at two distinct real points. If $\alpha_1, \ldots, \alpha_n$ satisfies equation (16.2), the direction of the hyperplane in the space $(t, x_1, \ldots, x_n)$, the normal to which is parallel to the vector $(1, \alpha_1, \ldots, \alpha_n)$, is characteristic (see Section 3).

We shall say that a surface K with a conical singular point for $t = t^0$, $x_i = x_i^0$ such that the hyperplane tangent to K at every point has a characteristic direction is a characteristic cone of equation (16.1) at the point $(t^0, x_1^0, x_2^0, \ldots, x_n^0)$

If
$$F(t, x_1, \ldots, x_n) = 0$$

is the equation of the surface of a characteristic cone (or in general of any characteristic surface (see Section 3)), the function F must satisfy the equation

$$\left(\frac{\partial F}{\partial t}\right)^2 = \sum_{i,j=1}^{n} A_{ij}\frac{\partial F}{\partial x_i}\frac{\partial F}{\partial x_j} + \sum_{i=1}^{n} A_{0i}\frac{\partial F}{\partial t}\frac{\partial F}{\partial x_i}.$$

At every point $(t^0, x_1^0, x_2^0, \ldots, x_n^0)$ of the region G in which equation (16.1) is t-hyperbolic, there will be located in this region a unique characteristic cone with vertex at that point, which intersects every hyperplane $t =$ const. along some closed surface S provided $|t - t^0|$ is sufficiently small. This cone and the portion of the hyperplane $t =$ const. that is bounded by the surface S bounds some region K'.

If $n = 1$, the characteristic cone degenerates into two lines l_1 and l_2 passing through the point (t^0, x_1^0); and the base of this cone degenerates into the segment of the straight line t — const. lying between the points of intersection of this straight line with the lines l_1 and l_2.

2. There exists a number L depending on n such that, for

all functions $\varphi_0(x_1, \ldots, x_n)$ and $\varphi_1(x_1, \ldots, x_n)$ possessing L continuous derivatives defined in some region G_0 of the hyperplane $t=t_0$, there exists exactly one continuous solution with continuous first and second derivatives of the t-hyperbolic equation (16.1) satisfying the conditions

$$\left.\begin{aligned} u(t_0, x_1, \ldots, x_n) &= \varphi_0(x_1, \ldots, x_n), \\ u'_t(t_0, x_1, \ldots, x_n) &= \varphi_1(x_1, \ldots, x_n). \end{aligned}\right\} \qquad (16.3)$$

This solution is uniquely determined by conditions (16.3) at every point $(t, x_1, \ldots, x_n)$ if the base of the characteristic cone with vertex at this point lies entirely in the region G_0. Let us denote by G the set of all such points $(t, x_1, \ldots, x_n)$.

If the functions $\varphi_0(x_1, \ldots, x_n)$ and $\varphi_1(x_1, \ldots, x_n)$ and their first L derivatives change by a sufficiently small amount, the corresponding solution to the Cauchy problem will also change only slightly throughout the entire region G. Thus, the Cauchy problem for equation (16.1) is stated correctly. $L=[n/2]+2$ for linear hyperbolic equations with constant coefficients containing only terms with second derivatives. S.L. Sobolev showed that $L \leqslant [n/2]+3$ for general linear second order equations. Here, it is assumed that the coefficients of the equation satisfy certain smoothness conditions, which are known to be satisfied when all the coefficients of the equation have continuous derivatives of order up to $[n/2]+2$ inclusively*.

3. We shall say that for equation (16.1) there is no wave diffusion in the region G in question of the space $(t, x_1, \ldots, x_n)$ if the solution u of the Cauchy problem at the vertex $(t, x_1, \ldots, x_n)$ of the characteristic cone depends only on the values of $\varphi_0(x_1, \ldots, x_n)$ and $\varphi_1(x_1, \ldots, x_n)$ and their derivatives on the boundary of the base of that cone for an arbitrary position of the characteristic within the region G. In the opposite case, we shall say that there is diffusion of waves. Hadamard † showed many years ago that there will always be diffusion of waves if n is either 1 or an even

* SOBOLEV, S. L., Nekotorye primeneniya funktsional'nogo analiza v matematicheskoy fizike *(Certain applications of functional analysis in mathematical physics)*, Leningrad (1950); *Matem. sbornik*, **1** (43); **1**, 39-72 (1936).

† HADAMARD, *Le problème de Cauchy*, Paris, 209-241 (1932).

number. Matisson investigated the case of $n=3$ in 1939. He found that for* $n=3$ all hyperbolic equations for which there is no diffusion of waves coincide, after certain minor transformations, with equation (13.1). All these equations are obtained from (13.1) by means of the following transformations:

(a) A change of independent variables,

(b) A linear change in the function u,

(c) Multiplication of both sides of the equation by some function of $t, x_1, \ldots, x_n$.

It has recently been shown that, for an arbitrary odd number $n \geqslant 5$, there exist hyperbolic equations for which there is no wave diffusion and which cannot be reduced to equation (13.1) by means of transformations of this nature**.

4. We have so far investigated in this section only the case in which the Cauchy conditions are given on the hyperplane $t=$ const. The case in which the Cauchy conditions are given on an arbitrary curved surface can be reduced to this particular case by a change of independent variables provided all the characteristic cones with vertices sufficiently close to this surface intersect it along closed $(n-1)$-dimensional surfaces.

5. The nonlinear equation

$$\frac{\partial^2 u}{\partial t^2} = F\left(t, x_1, \ldots, x_n, u, \frac{\partial u}{\partial t}, \ldots, \frac{\partial u}{\partial x_i}, \ldots \right.$$
$$\left. \ldots, \frac{\partial^2 u}{\partial x_i \partial x_j}, \ldots, \frac{\partial^2 u}{\partial t \partial x_j}, \ldots\right) \quad (16.4)$$

applying to a region G in the space $(t, x_1, \ldots, x_n)$ is said to be t-hyperbolic close to a certain function $u_0(t, x_1, \ldots, x_n)$ defined in the region G if the linear equation

$$\frac{\partial^2 u}{\partial t^2} = \sum_{\substack{i,j=1 \\ j \geq i}}^{n} A_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{j=1}^{n} A_{0j} \frac{\partial^2 u}{\partial t \partial x_j}, \quad (16.5)$$

where the A_{ij} are the partial derivatives of the right member

* MATISSON, *Acta Mathematica*, **71**, No. 3-4, 249 (1939).

STELLMACHER, *Math. Annalen*, **130, 3, 219-233 (1955).

of equation (16.4) with respect to $\frac{\partial^2 u}{\partial x_i \partial x_j}$ evaluated at

$$u \equiv u_0(t, x_1, \ldots, x_n)$$
$$(i = 0, 1, \ldots, n; \ j = 1, 2, \ldots, n; \ x_0 = t),$$

is t-hyperbolic in that region.

For the nonlinear equation (16.4), the Cauchy problem is correctly stated if, at the point $t = t_0$, conditions of the form

$$u(t_0, x_1, \ldots, x_n) = \varphi_0(x_1, \ldots, x_n),$$
$$u_t'(t_0, x_1, \ldots, x_n) = \varphi_1(x_1, \ldots, x_n),$$

are given such that equation (16.5) will be t-hyperbolic close to the function

$$u_0(t, x_1, \ldots, x_n) = \varphi_0(x_1, \ldots, x_n) + (t - t_0)\varphi_1(x_1, \ldots, x_n).$$

S.L. Sobolev showed* that $L \leqslant [n/2] + 4$ for a nonlinear hyperbolic equation. Here, it is assumed that the function F on the right-hand side of equation (16.4) has continuous derivatives with respect to all arguments of order $[n/2] + 3$.

6. The system of linear equations

$$\sum_{j=1}^{N} \sum_{k_0+k_1+\ldots+k_n \leqslant n_j} A_{ij}^{(k_0 k_1 \ldots k_n)}(t, x_1, \ldots, x_n) \frac{\partial^k u_j}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_n^{k_n}}$$
$$= f_i(t, x_1, \ldots, x_n) \quad (i = 1, 2, \ldots, N)$$

is said to be t-hyperbolic at the points $(t^0, x_1^0, \ldots, x_n^0)$ if, for arbitrary real values of α_i not all zero, the determinant

$$\left| \sum_{k_0+k_1+\ldots+k_n = n_j} A_{ij}^{(k_0 k_1 \ldots k_n)}(t^0, x_1^0, \ldots, x_n^0) \lambda^{k_0} \alpha_1^{k_1} \ldots \alpha_n^{k_n} \right|$$

has only distinct real roots λ. The definition for t-hyperbolicity of a nonlinear system close to any of its solutions is analogous.

It has been proven that the Cauchy problem is correctly

* SOBOLEV, S.L., *Dokl. Akad. Nauk*, **XX**, No. 2-3, 79-83 (1938); *Certain applications of functional analysis in mathematical physics*, Leningrad (1950).

stated for hyperbolic systems†.

For equations with constant coefficients, the definition of hyperbolicity was generalised by Gårding as follows: The equation

$$\sum_{0\leqslant k_0+k_1+\ldots+k_n\leqslant m} a_{k_0k_1\ldots k_n}\frac{\partial^k u}{\partial t^{k_0}\partial x_1^{k_1}\ldots\partial x_n^{k_n}}=0$$

is said to be hyperbolic relative to the direction $(\xi_0, \xi_1, \ldots, \xi_n)$, where the ξ_i are real and $\sum_{i=0}^{n}\xi_i^2>0$, if

$$\sum_{k_0+k_1+\ldots+k_n=m} a_{k_0k_1\ldots k_n}\xi_0^{k_0}\xi_1^{k_1}\ldots\xi_n^{k_n}\neq 0$$

and there exists a real number λ^* such that

$$\sum_{0\leqslant k_0+k_1+\ldots+k_n\leqslant m} a_{k_0k_1\ldots k_n}(\lambda\xi_0+i\alpha_0)^{k_0}(\lambda\xi_1+i\alpha_1)^{k_1}\cdot\ldots(\lambda\xi_n+i\alpha_n)^{k_n}\neq 0$$

for $\lambda>\lambda^*$ and arbitrary real values of α_i. It has been shown** that of all linear equations with constant coefficients, only for equations that are hyperbolic in the above sense is the Cauchy problem correctly stated for arbitrary sufficiently smooth initial functions given on the hyperplane

$$\xi_0x_0+\xi_1x_1+\ldots+\xi_nx_n=0^*$$

The application of Fourier transformations plays an important role in the study of equations with constant coefficients. By means of Fourier transformations, the question of the correctness of the statement of the Cauchy problem for systems of linear equations with constant coefficients or coefficients depending on t has

† PETROVSKII, I.G., *Matem. sbornik,* 2 (44), 815-870 (1937). See also LERAY, *Hyperbolic Differental Equations,* Princeton University Press (1953).

GÅRDING, *Acta Mathematica,* **85, No. 1-2, 1-62 (1951). See also GEL'FAND, I.M. and SHILOV, G.E., *Generalized functions,* **3**, Fizmatgiz, Chapter III (1958).

been studied and the qualitative properties of the solutions of such systems have been ascertained**.

7. For a t-hyperbolic equation with constant coefficients of the form

$$\sum_{k_0+k_1+\ldots+k_n=m} a_{k_0,\,k_1,\,\ldots,\,k_n} \frac{\partial^m u}{\partial t^{k_0}\partial x_1^{k_1}\ldots\partial x_n^{k_n}} = 0 \tag{16.6}$$

formulae have been obtained giving the solution to the Cauchy problem with initial conditions on the hyperplane $t=0$†.

The question of wave diffusion, which we have already studied for the wave equation, has been studied for equations of the form (16.6). Generally speaking, the lateral surface of the characteristic cone of equation (16.6) with vertex at the point $(t^*, \overset{*}{x}_1, \ldots, \overset{*}{x}_n)$ partitions its base on the hyperplane $t=0$ into several regions. We shall call one of these regions a lacuna if, for arbitrary (but sufficiently smooth) changes in the initial conditions that are made only within this region, the solution to the Cauchy problem for equation (16.6) does not change at the point $(t^*, \overset{*}{x}_1, \ldots, x_n)$. If the lacuna contains the projection of the vertex of the characteristic cone onto the hyperplane $t=0$, there will be no diffusion of waves for equation (16.6). The existence of lacunae for equation (16.6) is determined by the geometric (topological) properties of the surface

$$\sum_{k_0+k_1+\ldots+k_n=m} a_{k_0,\,k_1,\,\ldots,\,k_n} \lambda^{k_0} z_1^{k_1} \ldots z_n^{k_n} = 0$$

for $\lambda=1$ in the complex space $(z_1, z_2, \ldots, z_n)$. Necessary and sufficient conditions for the existence of lacunae have been found.

The question of wave diffusion and lacunae has also been studied for general t-hyperbolic systems‡.

PETROVSKII, I.G., *Moscow State University Bulletin*, Section A, **1, Issue 7 (1938); GEL'FAND, I.M. and SHILOV, G.E., *Generalized functions*, Issue 3, Fizmatgiz, Chapter III (1958).

† HERGLOTZ, *Berichte der Sächsischen Akademie*, **78**, 93-126, 287-318 (1926); **80**, 69-114 (1928). PETROVSKII, I.G., *Matem. sbornik*, **17**, 59, 3, 289-370 (1945). GEL'FAND, I.M. and SHAPIRO, Z.Ya., *Uspekhi matem. nauk*, **10**, 3, 3-70 (1955).

‡ PETROVSKII, I.G., *Izv. Akad. Nauk SSSR, seriya matem.*, **8**, 101-106 (1944); *Matem. sbornik*, **17** (59), 3, 289-370 (1945).

8. The following is an approximating method of solving the Cauchy problem (the method of finite differences) for the equation

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2} \tag{16.7}$$

under the initial conditions

$$u(0,\ x,\ y)=\varphi(x,\ y), \quad u_t'(0,\ x,\ y)=\psi(x,\ y).$$

Here, it is assumed that the initial functions $\varphi(x, y)$ and $\psi(x,\ y)$ have continuous derivatives up to the fourth order inclusively and that they are defined in some rectangle G:

$$a<x<b;\quad c<y<d.$$

Let us introduce three families of parallel planes in the space $(t,\ x,\ y)$:

$$t=k\Delta,\ k=0,\ 1,\ 2,\ 3,\ \dots,\qquad x=m\delta,\ y=n\delta.$$

Here, Δ and δ are positive numbers. The numbers m and n assume all positive integral values such that

$$a<m\delta<b \text{ and } c<n\delta<d.$$

To simplify the exposition, we assume that

$$a=m_1\delta,\ b=m_2\delta,\ c=n_1\delta,\ d=n_2\delta.$$

In equation (16.7), we replace $u_{tt}''(k\Delta,\ m\delta,\ n\delta)$ with

$$\frac{u[(k-1)\Delta,\ m\delta,\ n\delta]+u[(k+1)\Delta,\ m\delta,\ n\delta]-2u(k\Delta,\ m\delta,\ n\delta)}{\Delta^2},$$

$u_{xx}''(k\Delta,\ m\delta,\ n\delta)$ with

$$\frac{u[k\Delta,(m+1)\delta,\ n\delta]+u[k\Delta,(m-1)\delta,\ n\delta]-2u(k\Delta,\ m\delta,\ n\delta)}{\delta^2},$$

and $u_{yy}''(k\Delta,\ m\delta,\ n\delta)$ with

$$\frac{u[k\Delta,\ m\delta,\ (n+1)\delta]+u[k\Delta,\ m\delta,\ (n-1)\delta]-2u(k\Delta,\ m\delta,\ n\delta)}{\delta^2}.$$

It is easy to show that if $u(t,\ x,\ y)$ has continuous first and second derivatives, then, for sufficiently small Δ and

δ that result from such a substitution, the errors will be small. The differential equation (16.7) then becomes a difference equation, which we denote by (k, m, n). If we assign (k, m, n) various admissible values, we obtain a system of difference equations. We denote the solution of this system by $\bar{u}$.

Corresponding to the initial conditions, we obtain

$$\bar{u}(0, m\delta, n\delta) = \varphi(m\delta, n\delta),$$

$$\frac{\bar{u}(\Delta, m\delta, n\delta) - \bar{u}(0, m\delta, n\delta)}{\Delta} = \psi(m\delta, n\delta).$$

Then, the initial conditions define $\bar{u}(0, m\delta, n\delta)$ and $\bar{u}(\Delta, m\delta, n\delta)$ at all the nodes for which the corresponding points $(0, m\delta, n\delta)$ lie in the region G.

Then, when we write the difference equations $(1, m, n)$, we get values $\bar{u}(2\Delta, m\delta, n\delta)$ at all points $(2\Delta, m\delta, n\delta)$, that are apexes A' of pyramids of the type shown in Fig. 5. Here, it is assumed that all the points $[0, (m \pm 1)\delta, (n \pm 1)\delta]$ lie within the rectangle G, that is, that

$$m_1 + 1 < m < m_2 - 1, \quad n_1 + 1 < n < n_2 - 1.$$

If we then write the equations $(2, m, n)$, we find the values

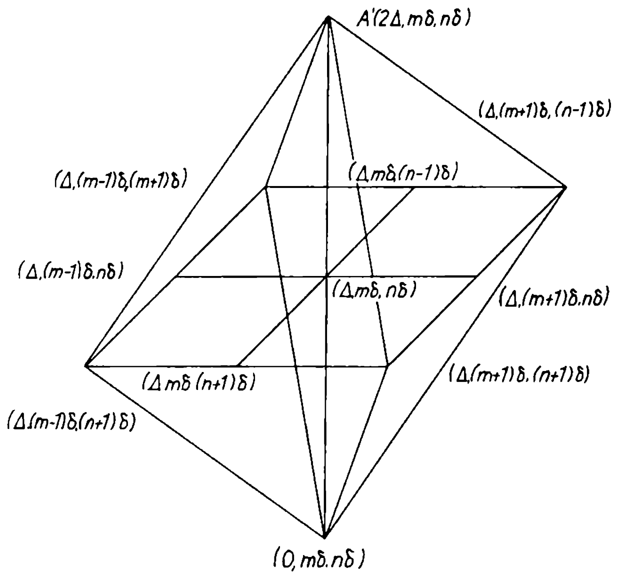


Fig. 5

of $\bar{u}$ at the points $(3\Delta, m\delta, n\delta)$, where

$$m_1+2<m<m_2-2, \quad n_1+2<n<n_2-2,$$

by using the values of $\bar{u}$ that are already found on the planes

$$t=\Delta, \quad t=2\Delta.$$

If we continue with these calculations, we find values of $\bar{u}$ at all points $(k\Delta, m\delta, n\delta)$, lying within a pyramid with base G on the plane $t=0$ and with lateral faces inclined at an angle $\arctan \frac{\Delta}{\delta}$ to this plane.

If $\Delta<\delta$ and δ is sufficiently small, it can be shown that the values found for $\bar{u}$ $(k\Delta, m\delta, n\delta)$ will differ by an arbitrarily small amount from the values at these points of the function $u(t, x, y)$ representing the exact solution of the posed Cauchy problem.

Approximate values of $u(t, x, y)$ for $t<0$ are determined analogously.

These constructions make it possible to obtain an approximate solution to the Cauchy problem for more general linear hyperbolic second-order equations with an arbitrary number of independent variables*.

Many books and papers have been devoted to the approximate solution of hyperbolic equations and systems by the method of finite differences.

9. The equation

$$\frac{\partial^2 u}{\partial y^2}=k(y)\,h(x, y)\frac{\partial^2 u}{\partial x^2}+a(x, y)\frac{\partial u}{\partial x}+b(x, y)\frac{\partial u}{\partial y}$$
$$+c(x, y)\,u+f(x, y), \tag{16.8}$$

where $k(0)=0$, $k(y)$ is a monotonic increasing function of y, and $h(x, y)>0$ for $y\geqslant 0$, is hyperbolic for $y>0$ and parabolic for $y=0$. It can be shown that the Cauchy problem for equation (16.8) with the initial conditions on the parabolic line $y=0$

$$u(x, 0)=\varphi(x), \quad \frac{\partial u}{\partial y}(x, 0)=\psi(x) \tag{16.9}$$

* See, for example, COURANT, FRIEDRICHS, LEWY, *Uspekhi matem. nauk,* Issue **VIII**, 147-160 (1941).

is correctly stated if the coefficients in the equation are sufficiently smooth functions and if

$$\lim_{y \to 0} \frac{ya(x, y)}{\sqrt{k(y)}} = 0.$$

It is possible to exhibit examples in which this condition is not satisfied but the Cauchy problem (16.8)-(16.9) is nonetheless correctly stated*. Analogous results are obtained also for equations with several independent variables.

10. Hyperbolic systems of nonlinear equations have a wide range of application in mechanics, especially in the study of gaseous motion. Many problems in mechanics lead to a consideration of discontinuous initial conditions and discontinuous solutions. The Cauchy problem for nonlinear hyperbolic systems of equations with discontinuous initial conditions has a number of peculiarities that linear systems of equations do not have. Let us consider some examples. As an initial condition for the Cauchy problem, let us take discontinuous functions of the form

$$\varphi(x) = \begin{cases} 1, & \text{if} \quad x \leqslant 0, \\ -1, & \text{if} \quad x > 0 \end{cases}$$

or

$$\psi(x) = \begin{cases} -1, & \text{if} \quad x \leqslant 0, \\ 1, & \text{if} \quad x > 0. \end{cases}$$

For the linear equation

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} = 0$$

the solution to the Cauchy problem with the initial condition

$$u(0, x) = \varphi(x) \tag{16.10}$$

and the solution with the initial condition

$$u(0, x) = \psi(x) \tag{16.11}$$

* BEREZIN, I.S., *Matem sbornik*, **24** (66): 2, 301-320 (1949). PROTTER, N.H., *Canadian Journal of Math.*, **6**: 4, 542-553 (1954).

are uniquely determined at all points of the half-plane $t > 0$. These solutions have a discontinuity* at points of the straight line $x - t = 0$.

For the nonlinear equation

$$u\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} = 0 \tag{16.12}$$

the solutions to the Cauchy problem with initial conditions (16.10) and (16.11) are not uniquely determined even in an arbitrarily small neighbourhood of the straight line $t = 0$ on which the initial conditions are given.

To see this, let us draw the characteristics of equation (16.12) through the points $(0, x, u(0, x))$ in the space (t, x, u). These characteristics are straight lines parallel to the tx-plane**. If $u(0, x) = \varphi(x)$, the projections of these characteristics onto the tx-plane cover all points of the half-plane $t > 0$. Points of the region Q between the straight lines $x - t = 0$ and $x + t = 0$ are twice covered by the projections of these characteristics and, what is more, with

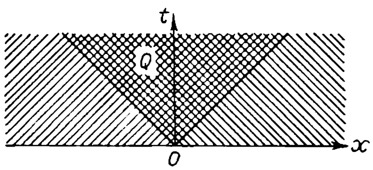


Fig. 6

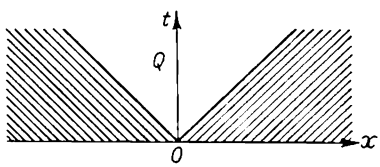


Fig. 7

different values of u (see Fig. 6). From this, it is easy to see that the solution $u(t, x)$ at points between the straight lines $x - t = 0$ and $x + t = 0$ cannot be uniquely determined from the initial conditions.

If $u(0, x) = \psi(x)$, the projections of the characteristics of

* Discontinuous solutions of linear hyperbolic systems are studied in the paper by COURANT and LAX, *Proc. Nat. Acad. Sci., USA*, **42**: 11, 872-876 (1956).

** The characteristic of equation (16.12) that passes through the point $(0, x_0, u(0, x_0))$ is given by the equations

$$u = u(0, x_0) \text{ and } x = u(0, x_0)\, t + x_0.$$

See PETROVSKII, I.G., *Op. cit.*, Section 55.

equation (16.12) passing through the points $(0, x, \psi(x))$ in the space (t, x, u) cover only points that do not belong to the region Q (see Fig. 7); that is, the solution cannot be determined from the initial condition at points lying between the straight lines $x-t=0$ and $x+t=0$.

Thus, for a unique determination of the solution to the Cauchy problem in the half-plane $t>0$ for the nonlinear equation (16.12) with initial conditions (16.10) or (16.11), the Cauchy problem needs to be restated.

Thus, for a hyperbolic system of equations describing the one-dimensional motion of a gas, supplementary relationships between the desired functions are introduced on the lines of discontinuity. This hyperbolic system was studied by Riemann. However, not all the supplementary conditions on the lines of discontinuity indicated by Riemann are satisfied in the case of real physical processes. Relationships on the lines of discontinuity for this hyperbolic system were correctly given by Hugoniot*. These relationships can be obtained by solving the system of equations describing the motion of a gas with due consideration of the viscosity and thermal conductivity and letting the coefficients of viscosity and thermal conductivity approach 0. This consideration of the viscosity and thermal conductivity results in introducing into a first-order system of equations second-order derivatives containing a small parameter as a coefficient.

We may define the solution to the Cauchy problem for equation (16.12) with an initial condition at $t=0$ as the limit as ε approaches zero of the solution of the equation

$$\varepsilon \frac{\partial^2 u}{\partial x^2} = u \frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} \quad (\varepsilon > 0)$$

with the same initial condition at $t=0$. In this case, the solution of the Cauchy problem may be a discontinuous function. On the line of discontinuity of the solution to equation (16.12), the following conditions will be satisfied:

$$u(t, x+0) < u(t, x-0) \quad (t>0)$$

$$\frac{dx}{dt} = \frac{u(t, x+0) + u(t, x-0)}{2},$$

* LANDAU, L.D. and LIFSHITZ, E.M., *The Mechanics of Continuous Media* (1954).

where $\frac{dx}{dt}$ is the (trigonometric) tangent of the angle between the (geometric) tangent to the line of discontinuity and the t-axis and where $u(x+0)$ and $u(x-0)$ denote respectively the right- and left-hand limits of the function $u(x)$ at the point x.

The function $u(t, x)$ defined by

$$u(t, x) = \begin{cases} 1 & \text{for } x<0 \\ -1 & \text{for } x>0 \end{cases}$$

is the solution to the Cauchy problem in its new statement for equation (16.12) with initial conditions (16.10).

The function $u(t, x)$ defined by

$$u(t, x) = \begin{cases} 1 & \text{for } x-t>0 \\ -1 & \text{for } x+t<0 \end{cases}$$

is a solution to the Cauchy problem in its new statement with initial condition (16.11). At every point of the half-plane $t>0$ lying in the region Q (that is, between the straight lines $x-t=0$ and $x+t=0$), the function $u(t, x)$ is equal to the tangent of the angle of inclination of the straight line connecting the point in question with the coordinate origin to the t-axis. The positions of the projections of the characteristics lying on this solution are shown in Fig. 8.

This function $u(t, x)$ is discontinuous for $t>0$. It is interesting to note that in this statement of the Cauchy problem there may be a continuous solution even though there is a discontinuity in the initial condition.

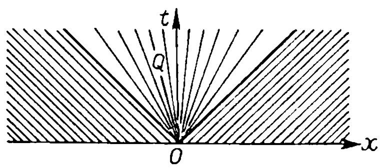


Fig. 8

In the case of smooth initial conditions, for linear hyperbolic equations, a smooth solution will be determined by initial conditions at all points of the half-plane $t>0$ for all initial conditions if the coefficients are subjected to certain

restrictions. For nonlinear hyperbolic equations, a smooth solution usually exists only in a small neighbourhood of the line on which the initial conditions are given. This fact makes it necessary to examine discontinuous solutions of non-linear hyperbolic equations.

The basic problem in the study of discontinuous solutions of nonlinear hyperbolic systems consists in determining the class of functions in which there exists a unique generalised solution of the Cauchy problem that depends continuously in some specified sense on the initial conditions. This question has been thoroughly studied for the general quasi-linear first-order equation*. It turns out that the qualitative properties of the generalised solutions of such an equation call to mind the properties of the solutions of the system of equations of gas dynamics. For this reason, the simplest quasilinear first-order equation

$$\frac{\partial u}{\partial t}+\frac{\partial \varphi(u)}{\partial x}=0$$

is often called the model equation of gas dynamics.

The question of discontinuous solutions of nonlinear hyper-bolic systems has as yet been only slightly investigated**. This question is of great theoretical interest and is very important in its applications.

* OLEINIK, O.A., *Uspekhi matem. nauk,* **12**, 3, 3-73, (1957): See also *Uspekhi matem. nauk,* **14**, 2, 159-170 (1959).

GEL'FAND in an article in *Uspekhi matem. nauk,* **14, 2, 87-158 (1959) reviews some of the results obtained in connection with this problem and discusses the formulation of a number of problems.

PART 2

OSCILLATIONS IN BOUNDED BODIES

17. INTRODUCTION

1. The preceding part of Chapter II was devoted to the Cauchy problem. Our attention was primarily directed to the wave equation (13.1), which the vibrations of homogeneous isotropic elastic bodies obey. The study of the function $u(t, x_1, \ldots, x_n)$, characterising these vibrations at the points $(x_1, x_2, \ldots, x_n)$ for t sufficiently close to the initial instant is reduced to the Cauchy problem. Since the value of the solution $u(t, x_1, \ldots, x_n)$ of equation (13.1) at the vertex $P(t, x_1, \ldots, x_n)$ of the characteristic cone is completely determined by the values of the initial functions φ_0 and φ_1 on the base C_p of that cone, when we study the function $u(t, x_1, \ldots, x_n)$ we may neglect the influence of the boundary once C_p leaves the region in which the functions φ_0 and φ_1 are given, that is, once C_p no longer intersects the boundary of the body. In this sense, in the preceding part of the chapter we were studying the vibrations of infinite or unbounded bodies.

In this part, we shall include the effect of these boundaries on the vibrations. Still confining ourselves to vibrations of homogeneous isotropic bodies, we arrive at the problem of finding solutions of the equation

$$\frac{\partial^2 u}{\partial t^2} = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}, \tag{17.1}$$

satisfying the initial conditions at $t=0$

$$\left.\begin{aligned} u(0, x_1, \ldots, x_n) &= \varphi_0(x_1, \ldots, x_n), \\ u_t'(0, x_1, \ldots, x_n) &= \varphi_1(x_1, \ldots, x_n), \end{aligned}\right\} \tag{17.2}$$

when the point $(x_1, \ldots, x_n)$ belongs to the given region G, and the boundary conditions given for all values of t on the boundary G. We shall consider only homogeneous boundary conditions of the form

$$u=0, \tag*{$(17.3)_1$}$$

$$\frac{\partial u}{\partial n}=0, \tag*{$(17.3)_2$}$$

$$\frac{\partial u}{\partial n}+\sigma u=0, \tag*{$(17.3)_3$}$$

where σ is some nonnegative continuous function independent of t that is given on the boundary G and $\frac{\partial}{\partial n}$ denotes differentiation along the direction of the outer normal to the surface G (see Section 1).

Some physical problems, for example, the problems of the vibrations of nonhomogeneous elastic bodies, reduce to finding the solutions of equations of the form

$$\rho\frac{\partial^2 u}{\partial t^2}=\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\left(p_i\frac{\partial u}{\partial x_i}\right)-qu+f. \tag{17.4}$$

under the same boundary conditions (17.3) and initial conditions (17.2). Here, ρ, p_i, q and f are sufficiently smooth functions of x_k. Ordinarily,

$$\rho>\rho_0>0; \quad p_i>p_{i0}>0; \quad q\geqslant 0.$$

Since the wave equation (17.1) and equation (17.4) do not change if we replace t with $-t$, our reasoning with regard to the solution of these equations for $t>0$ will also hold for $t<0$.

The problem of finding the solution to equation (17.4) under the initial conditions (17.2) and one of the boundary conditions (17.3) is called the mixed problem. The entire present part of Chapter II will be devoted to this problem.

2. The mixed problem is not the only possible problem for equation (17.1) or (17.4) in a bounded region. Practical problems often lead to other problems for these equations. We shall exhibit a number of such problems for the very simple equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}. \tag{17.5}$$

(1) Goursat's problem. Find the solution of equation (17.5) from its values on two segments of the characteristics.

On the segment OA (see Fig. 9) of the characteristic $t + x = 0$,

$$u(t, x) = \varphi(x).$$

On the segment OB of the characteristic $t - x = 0$,

$$u(t, x) = \psi(x).$$

For the solution to be continuous under these conditions, it is necessary that

$$\varphi(0) = \psi(0).$$

(See S.L. Sobolev, Uraveniya matematicheskoi fiziki (Equations of mathematical physics), Gostekhizdat, 1954, pp. 63-67.)

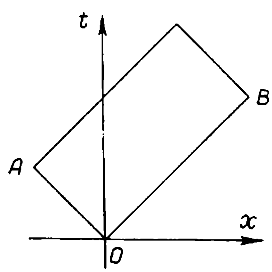


Fig. 9

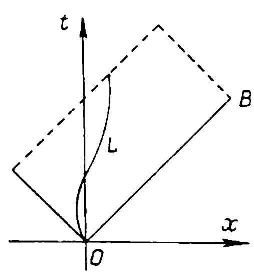


Fig. 10

(2) Find the solution to equation (17.5) if its values are given on the segment OB of the characteristic $t = x$ and on a curve L passing through the point O that lies within the angle formed by the characteristics $t = \pm x$ and that possesses the property that every characteristic $t = x + C$ intersects L at one point (see Fig. 10).

The reader may solve these problems himself by using the representation of the solution (17.5) in the form

$$u(t, x) = f_1(t+x) + f_2(t-x)$$

(see Example 1 of Section 6).

In both cases, the solution is determined in the rectangle formed by the characteristics passing through the ends of the lines on which the values of the function u are given.

(3) If the values of the function $u(t, x)$ are given on two lines L and L_1 (which for simplicity we shall assume to be straight) that pass through the coordinate origin, there will be two quite different cases: (a) when L and L_1 lie within an angle formed by the characteristics passing through the origin and (b) when L and L_1 are separated by a characteristic.

In the first case, to determine uniquely the solution of equation (17.5), it is sufficient to give only the values of the function itself $u(t, x)$ on the lines L and L_1. In the second case, it is necessary to give the 'Cauchy' conditions, that is, the values of the solution itself on one of these lines and its first derivative along the normal to that line (see Goursat: A Course in Mathematical Analysis).

3. In the majority of cases that follow, our remarks will be equally applicable for any value of n. For greater convenience in calculations and drawings, we shall discuss only the case in which n is 2 or 1 and shall make explicit statements for other values of n only when the situation is considerably different from the case of $n=2$ or 1.

Assuming therefore that $n=2$, we shall examine solutions $u(t, x_1, x_2)$ of equations of the type (17.1) or (17.4) for $0 \leqslant t \leqslant T$ when the point (x_1, x_2) falls within a region G bounded by a curve l consisting of a finite number of arcs l_i with a continuously turning tangent.

In other words, taking $n=2$, we shall examine solutions $u(t, x_1, x_2)$ of equations (17.1) or (17.4) defined within a cylinder C_T whose generators are parallel to the t-axis and pass through the boundary of the region G in the plane $t=0$ and whose bases lie in the planes $t=0$ and $t=T$. Throughout this part of Chapter II, we shall always assume without always explicitly mentioning the fact that the solutions $u(t, x_1, x_2)$ in question satisfy equation (17.1) or (17.4) within

C_T and that they and their first two derivatives are continuous in $\bar{C}_T$, that is, in the cylinder C_T and on its boundary.

18. UNIQUENESS OF THE SOLUTION OF THE MIXED PROBLEM

Suppose that $u_1(t, x_1, x_2)$ and $u_2(t, x_1, x_2)$ are two solutions of the equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}, \tag{18.1}$$

that they are defined in a cylinder C_T, that they possess all the properties enumerated in the preceding section, and that they are solutions of the same mixed problem; that is, we shall assume that for $t=0$ they satisfy the same initial conditions (17.2) and that on the surface of the cylinder they satisfy the same boundary conditions of one of the forms (17.3). Our intention is to show that the functions $u_1(t, x_1, x_2)$ and $u_2(t, x_1, x_2)$ coincide everywhere in the cylinder C_T. Proof of this assertion is equivalent to proof of the following theorem.

Theorem. Suppose that the function

$$u(t, x_1, x_2) = u_2(t, x_1, x_2) - u_1(t, x_1, x_2),$$

satisfies equation (18.1) inside C_T, that it and its first and second derivatives are continuous within $\bar{C}_T$, that it satisfies one of the conditions (17.3) on the lateral surface of C_T, and that u and u_t vanish everywhere within the cylinder at $t=0$. Then, u vanishes identically in C_T.

Proof: Consider the integral

$$\iiint \frac{\partial u}{\partial t}\left(\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2}\right) dt\, dx_1\, dx_2, \tag{18.2}$$

over the cylinder C_{t^*}, where $0 < t^* \leqslant T$. Since the function u satisfies equation (18.1), the integral is equal to zero. Let us transform it into an integral over the surface of the cylinder C_{t^*} in a manner analogous to what was done in Section 11 in the proof of the uniqueness of the solution to the Cauchy problem. We obtain

$$\frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]_{t=t^*} dx_1\,dx_2$$
$$-\frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]_{t=0} dx_1\,dx_2$$
$$-\int\limits_0^{t^*} dt \int\limits_l \left[\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_1}\cos(n, x_1)+\frac{\partial u}{\partial t}\frac{\partial u}{\partial x_2}\cos(n, x_2)\right] ds = 0.$$

Here, l, as usual, denotes the boundary of the region G and ds is an element of arc of the boundary. The first integral is taken over the upper base of the cylinder C_{t^*} and the second over the lower base. The third integral is taken over the lateral surface. The last integral can be rewritten in the form

$$\int\limits_0^{t^*} dt \int\limits_l \frac{\partial u}{\partial t}\frac{\partial u}{\partial n}\,ds.$$

Thus, we finally obtain

$$\frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]_{t=t^*} dx_1\,dx_2$$
$$-\frac{1}{2}\iint\limits_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]_{t=0} dx_1\,dx_2$$
$$-\int\limits_0^{t^*} dt \int\limits_l \frac{\partial u}{\partial t}\frac{\partial u}{\partial n}\,ds = 0. \tag{18.3}$$

The second of these integrals is equal to zero by virtue of the initial conditions. If $u=0$ everywhere on the boundary of G, the derivative $\frac{\partial u}{\partial t}$ must be equal to zero on that boundary. Therefore, the third integral must be equal to zero. It is also equal to zero in the case in which $\frac{\partial u}{\partial n}=0$ on the boundary. On the other hand, if $\frac{\partial u}{\partial n}+\sigma u=0$, on the boundary, this integral becomes

$$\begin{aligned}-\int\limits_0^{t^*} dt \int\limits_l \sigma u\,\frac{\partial u}{\partial t}\,ds &= -\frac{1}{2}\int\limits_l \sigma\,ds \int\limits_0^{t^*} \frac{\partial (u^2)}{\partial t}\,dt \\ &= -\frac{1}{2}\int\limits_l \sigma u^2(t^*)\,ds+\frac{1}{2}\int\limits_l \sigma u^2(0)\,ds.\end{aligned} \tag{18.4}$$

The last integral is equal to zero because of the initial conditions. Thus, for every t^* between 0 and t, we have

$$\iint_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]_{t=t^*} dx_1\,dx_2=0, \tag{18.5}$$

if either $u=0$ or $\frac{\partial u}{\partial n}=0$ everywhere on the boundary G, and we have

$$\iint_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]_{t=t^*} dx_1\,dx_2 + \frac{1}{2}\int_l \sigma u^2(t^*)\,ds=0 \tag{18.6}$$

if $\frac{\partial u}{\partial n}+\sigma u=0$, everywhere on the boundary of G Since $\sigma \geqslant 0$, it follows from the relations (18.5) and (18.6) that, under any of the boundary conditions (17.3),

$$\iint_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\left(\frac{\partial u}{\partial x_2}\right)^2\right]_{t=t^*} dx_1\,dx_2=0. \tag{18.7}$$

Since we assume that all the first derivatives of the function u are continuous in $\bar{C}_T$ and since t^* is an arbitrary number between 0 and t, it follows from equation (18.7) that

$$\frac{\partial u}{\partial t}=\frac{\partial u}{\partial x_1}=\frac{\partial u}{\partial x_2}=0$$

everywhere in $\bar{C}_T$. This means that u is constant throughout C_T. And since $u(0, x_1, x_2)\equiv 0$, we have

$$u(t, x_1, x_2)\equiv 0,$$

throughout the cylinder C_T, which completes the proof.

We note that the integral on the left side of (18.7) is equal, except for a constant factor, to the sum of the kinetic and potential energies of a vibrating membrane at the instant $t=t^*$. Equation (18.3) under the boundary conditions $(17.3)_1$ and $(17.3)_2$ expresses the law of conservation of energy (see Section 1, subsection 3).

Problem. Prove the uniqueness in C_T of the solution to

the problem with initial conditions (17.2) and boundary conditions $(17.3)_1$ for equation (17.4).

19. CONTINUOUS DEPENDENCE OF THE SOLUTION ON THE INITIAL CONDITIONS

Theorem. Suppose that we have two functions $u_1(t, x)$ and $u_2(t, x)$ satisfying equation (17.1) in the cylinder C_T with $n=1$*. Suppose that both these solutions satisfy the same boundary conditions of one of the types shown in (17.3) on the lateral surface and that, at $t=0$,

$$u_1(0, x)=\varphi_0^{(1)}(x); \qquad u'_{1t}(0, x)=\varphi_1^{(1)}(x);$$
$$u_2(0, x)=\varphi_0^{(2)}(x); \qquad u'_{2t}(0, x)=\varphi_1^{(2)}(x).$$

If the differences

$$\varphi_i^{(2)}(x)-\varphi_i^{(1)}(x)=\varphi_i(x) \qquad (i=0,1)$$

and the first derivative of the function $\varphi_0(x)$ are sufficiently small in absolute value throughout G, the difference

$$u_2(t, x)-u_1(t, x)=u(t, x)$$

will be arbitrarily small in absolute value throughout C_T.

An analogous theorem is valid for solutions of equation (17.1) in C_T for arbitrary n. But then, to ensure that

$$u_2(t, x_1, \ldots, x_n)-u_1(t, x_1, \ldots, x_n)=u(t, x_1, \ldots, x_n)$$

will be small throughout the entire cylinder C_T, we must require that not only the functions

$$u(0, x_1, \ldots, x_n) \text{ and } u_t(0, x_1, \ldots, x_n),$$

but also their first $[n/2]+1$ derivatives with respect to $x_1, \ldots, x_n$ differ only slightly from zero to G. Furthermore, it is necessary that the first $[n/2]$ derivatives of these differences satisfy certain supplementary conditions on the boundary of the region G constituting the base of the cylinder C_T that are automatically satisfied for $n=1$.

* Obviously, for $n=1$ the cylinder C_T is a rectangle with sides parallel to the t- and x-axes.

The proof of this theorem for $n > 1$ is much more complicated than for the case of $n=1$, and we shall omit it.

Proof of the theorem for $n=1$. Let us again consider an integral of the type (18.2) over a cylinder C_T, which has now degenerated into the rectangle $\{0 \leqslant t \leqslant T, \quad a \leqslant x \leqslant b\}$. As before, this integral is equal to zero for every t^* between 0 and T. By a transformation analogous to the one made above, we obtain

$$\iint_{\mathcal{U}_{t^*}} \frac{\partial u}{\partial t}\left(\frac{\partial^2 u}{\partial t^2}-\frac{\partial^2 u}{\partial x^2}\right) dt\,dx = \frac{1}{2}\int_a^b \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x}\right)^2\right]_{t=t^*} dx$$
$$-\frac{1}{2}\int_a^b \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x}\right)^2\right]_{t=0} dx -$$
$$-\int_0^{t^*}\left(\frac{\partial u}{\partial t}\frac{\partial u}{\partial x}\right)_{x=b} dt+\int_0^{t^*}\left(\frac{\partial u}{\partial t}\frac{\partial u}{\partial x}\right)_{x=a} dt=0.$$

Since $a<b$ and†

$$u(0, x)=\varphi_0; \quad u_t'(0, x)=\varphi_1; \quad \left.\frac{\partial u}{\partial x}\right|_{x=b}=\frac{\partial u}{\partial n}; \quad \left.\frac{\partial u}{\partial x}\right|_{x=a}=-\frac{\partial u}{\partial n}$$

we then have

$$\frac{1}{2}\int_a^b \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x}\right)^2\right]_{t=t^*} dx=\frac{1}{2}\int_a^b [\varphi_1^2(x)+\varphi_0'^2(x)]\,dx$$
$$+\int_0^{t^*}\left.\frac{\partial u}{\partial t}\frac{\partial u}{\partial n}\right|_{x=a} dt+\int_0^{t^*}\left.\frac{\partial u}{\partial t}\frac{\partial u}{\partial n}\right|_{x=b} dt. \tag{19.1}$$

If $u(t, a)=0$ or if $\frac{\partial u(t, a)}{\partial n}=0$, then

$$\int_0^{t^*}\left.\frac{\partial u}{\partial t}\frac{\partial u}{\partial n}\right|_{x=a} dt=0.$$

On the other hand, if at $x=a$ the boundary condition $\frac{\partial u}{\partial n}+\sigma_a n=0$ is satisfied, we have

† We recall that $\partial/\partial n$ always denotes differentiation along the outwardly-directed normal.

$$\int_0^{t^*} \frac{\partial u}{\partial t}\frac{\partial u}{\partial n}\bigg|_{x=a} dt = -\sigma_a \int_0^{t^*} u\frac{\partial u}{\partial t}\,dt = -\frac{\sigma_a u^2(t^*, a)}{2} + \frac{\sigma_a u^2(0, a)}{2}.$$

Analogous equations may be written for $x=b$ if at $x=b$ one of the conditions $u(t, b)=0$, $\frac{\partial u(t, b)}{\partial n}=0$, $\frac{\partial u}{\partial n}+\sigma_b u=0$ is imposed.

Thus, by eliminating the negative terms on the right-hand side of formula (19.1) if necessary, we have, in the case of each of the boundary conditions (17.3),†

$$\int_a^b \left[\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial u}{\partial x}\right)^2\right]_{t=t^*} dx \leqslant \int_a^b [\varphi_1^2(x) + \varphi_0'^{\,2}(x)]\,dx + \sigma_a\varphi_0^2(a) + \sigma_b\varphi_0^2(b) \tag{19.2}$$

Since the right-hand side is assumed to be small, the left-hand side is also small. Denoting by ε^2 the value of the right-hand side of inequality (19.2), we see, that, for every t^* between 0 and T and for every x such that $a \leqslant x \leqslant b$,

$$\int_a^x \left(\frac{\partial u}{\partial x}\right)^2_{t=t^*} dx \leqslant \varepsilon^2, \tag{19.3$_1$}$$

$$\int_a^x \left(\frac{\partial u}{\partial t}\right)^2_{t=t^*} dx \leqslant \varepsilon^2. \tag{19.3$_2$}$$

From inequality $(19.3)_1$, we obtain by use of Bunyakovskii's inequality,

$$|u(t^*, x) - u(t^*, a)| \leqslant \int_a^x \left|\frac{\partial u}{\partial x}\right| dx$$

$$= \int_a^x 1\left|\frac{\partial u}{\partial x}\right| dx \leqslant \left[\int_a^x dx \int_a^x \left(\frac{\partial u}{\partial x}\right)^2 dx\right]^{\frac{1}{2}} \leqslant \sqrt{b-a}\,\varepsilon. \tag{19.4}$$

In just the same way, we obtain from inequality $(19.3)_2$

† Under the boundary conditions $u=0$ or $\partial u/\partial n = 0$, this inequality becomes an equality, expressing the law of conservation of energy.

$$\left|\frac{\partial}{\partial t}\int_a^b u\,dx\right| = \left|\int_a^b \frac{\partial u}{\partial t}\,dx\right| \leqslant \int_a^b \left|\frac{\partial u}{\partial t}\right| dx \leqslant \sqrt{b-a}\,\varepsilon. \tag{19.5}$$

Furthermore,

$$\left|\int_a^b u(t^*, x)\,dx - \int_a^b u(0, x)\,dx\right|$$

$$= \left|\int_0^{t^*}\left[\frac{\partial}{\partial t}\int_a^b u(t, x)\,dx\right]dt\right| \leqslant t^*\varepsilon\sqrt{b-a}.$$

From this we obtain

$$\left|\int_a^b u(t^*, x)\,dx\right| \leqslant t^*\varepsilon\sqrt{b-a} + \left|\int_a^b \varphi_0(x)\,dx\right|$$

$$\leqslant t^*\varepsilon\sqrt{b-a} + \max|\varphi_0|(b-a). \tag{19.6}$$

Integrating inequality (19.4) with respect to x from a to b we obtain

$$\left|\int_a^b u(t^*, x)\,dx - (b-a)\,u(t^*, a)\right| \leqslant \varepsilon(b-a)^{\frac{3}{2}} \tag{19.7}$$

If we use (19.6) to get an upper bound for the integral on the left-hand side of formula (19.7), we see that the values of $|u(t^*, a)|$ can be made arbitrarily small for every t^* in the interval $(0, T)$ and for sufficiently small ε and max $|\varphi_0|$. By again applying inequality (19.4), we see that $|u(t^*, x)|$ is small throughout the entire rectangle C_T, which completes the proof.

Remark 1: Suppose that the condition $u=0$ for all nonnegative values of $t \geqslant 0$ is given at one end point of the interval $[a, b]$, let us say at $x=a$. Then, the smallness of $|u(t, x)|$ through C_T follows immediately from the relation (19.4).

If the condition

$$\frac{\partial u}{\partial n} + \sigma_a u = 0$$

for

$$\sigma_a > 0$$

were given at one end point of the interval $[a, b]$, let us

say at $x=a$, the relation (19.1) would imply that

$$\int_a^b \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x}\right)^2\right]_{t=t^*} dx+\sigma_a u^2(t^*, a)$$

$$\leqslant \int_a^b [\varphi_1^2(x)+\varphi_0'^2(x)]\,dx+\sigma_a\varphi_0^2(a)+\sigma_b\varphi_0^2(b)\leqslant \varepsilon^2.$$

Therefore,

$$\sigma_a u^2(t^*, a)\leqslant \varepsilon^2,$$

and the smallness of $|u(t^*, x)|$ throughout C_T would again follow immediately from (19.4) because of the smallness of $|u(t^*, a)|$.

Remark 2: Consider the case of n greater than 1. Then just as when $n=1$ we would see that the uniform smallness of $|u_t'(0, x_1, \ldots, x_n)|$, $|u(0, x_1, \ldots, x_n)|$ and the derivatives

$$\left|\frac{\partial u(0, x_1, \ldots, x_n)}{\partial x_i}\right|$$

would imply that

$$\int_G \left[\left(\frac{\partial u}{\partial t}\right)^2+\left(\frac{\partial u}{\partial x_1}\right)^2+\cdots+\left(\frac{\partial u}{\partial x_n}\right)^2\right]_{t=t^*} dx_1\ldots dx_n. \qquad (19.8)$$

is small. But smallness of this integral for all t^* between 0 and T and smallness of $|u(0, x_1, \ldots, x_n)$ do not imply that $|u|$ is small throughout C_T. It is possible to exhibit a function u, for which this integral is small for all values of t^* in question but that nonetheless assumes quite large values at certain points of C_T despite the smallness of $|u(0, x_1, \ldots, x_n)|$. To ensure that $|u|$ will be small throughout C_T, it will be sufficient if, besides smallness of the integral (19.8), the following conditions are satisfied: At $t=t^*$, all integrals of the form (19.8), in which instead of u we have any possible derivative of the form

$$\frac{\partial^k u}{\partial t^{k_0}\partial x_1^{k_1}\ldots\partial x_n^{k_n}} \quad \text{for } k\leqslant\left[\frac{n}{2}\right],$$

are small and the values of $|u(0, x_1, \ldots, x_n)|$ are uniformly

small*. In just the same way, we can show that $|u|$ will be small throughout C_T if φ_0, φ_1, and their first $[n/2]+1$ derivatives with respect to $x_1, \ldots, x_n$ are sufficiently small in absolute value and if certain additional conditions regarding the values of φ_0 and φ_1 are satisfied on the boundary G**.

Remark 3: It is clear from the proof of the theorem that the conclusion of the theorem remains valid if we replace the requirement of uniform smallness of $|u(0, x)|$, $|u'_x(0, x)|$ and $|u_t(0, x)$ with the requirement of smallness of the integrals

$$\int_a^b u_x'^2(0, x)\,dx \text{ and } \int_a^b u_t'^2(0, x)\,dx$$

and one of the quantities $|u(0, a)|$ or $|u(0, b)|$. This is true because in inequalities (19.2) and (19.6) we used only the smallness of the integrals and the smallness of $|u(0, x)|$. But if $|u(0, a)|$ is small, then

$$|u(0, x)| = |u(0, a) + \int_a^x u'_x(0, x)\,dx|$$
$$\leqslant |u(0, a)| + \sqrt{b-a}\left(\int_a^b u_x'^2(0, x)\,dx\right)^{\frac{1}{2}},$$

which implies uniform smallness of $|u(0, x)|$.

Problem 1. Prove the theorem on the continuous dependence of the solution on the initial conditions for equation (17.4) with $n=1$ under the boundary condition $(17.3)_1$.

2. Show that the solution of the equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(p(x)\frac{\partial u}{\partial x}\right) - qu + f(t, x)$$

(where $p(x) > 0$, $q > 0$ and $f(t, x)$ are sufficiently smooth functions) satisfying the initial conditions (17.2) and the boundary condition $(17.3)_1$ changes by an arbitrarily small amount in C_T if the change in the function $f(t, x)$ is sufficiently small in C_T.

* This follows from the so-called embedding theorems of Sobolev (see SOBOLEV, S.L., *Certain Applications of functional analysis to mathematical physics,* Leningrad (1950).

See KRZYZANSKI, M. and SCHAUDER, J., *Studia Mathematica,* **VI, 162-189 (1936).

20. THE FOURIER METHOD FOR THE STRING EQUATION

1. In many cases, the so-called Fourier method can be applied to the solution of the mixed problem. In the present section, we shall examine the application of this method to a single particular example. We shall give the general outline of the application of this method to the solution of the mixed problem for a linear second-order equation with two independent variables.

Suppose that we wish to find the solution of the equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \tag{20.1}$$

that satisfies the initial conditions

$$\left.\begin{aligned} u(0, x) &= \varphi_0(x), \\ u'_t(0, x) &= \varphi_1(x), \\ 0 \leqslant x &\leqslant l, \end{aligned}\right\} \tag{20.2}$$

and the boundary conditions for $t > 0$

$$u(t, 0) = u(t, l) = 0. \tag{20.3}$$

First, let us seek the nontrivial (that is, not identically equal to zero) solutions of equation (20.1) of the form

$$u(t, x) = T(t) X(x), \tag{20.4}$$

that satisfy the boundary conditions (20.3). Here, we assume that $T(t)$ depends only on t and $X(x)$ depends only on x. When we substitute the right member of (20.4) in place of u in equation (20.1), we obtain

$$XT'' = X''T \quad \text{or} \quad \frac{T''}{T} = \frac{X''}{X}. \tag{20.5}$$

The left member of the second of these equations is independent of x and the right member is independent of t. Consequently, each of the quantities $\frac{T''}{T}$ and $\frac{X''}{X}$ is independent both of x and of t; in other words, it is constant. If we denote this constant value by λ, we have, from (20.5),

$$T'' + \lambda T = 0, \tag{20.6}$$

$$X'' + \lambda X = 0. \tag{20.7}$$

Thus, we have broken equation (20.5) down into two equations, one of which contains only functions of t and the other only functions of x. In such cases, we say that we have separated the variables.

To obtain a nontrivial solution $u(t, x)$ of the form (20.4) satisfying the boundary conditions (20.3), we need to find a nontrivial (that is, not identically equal to zero) solution of equation (20.7) satisfying the boundary conditions

$$X(0) = X(l) = 0. \tag{20.8}$$

The formulae giving the general solution of equation (20.7) have quite different forms depending on whether

$$\lambda < 0, \ \lambda = 0 \ \text{ or } \ \lambda > 0.$$

Let us consider each of these three cases.

Case I ($\lambda < 0$). Then the general solution of equation (20.7) can be written in the form

$$X(x) = C_1 e^{\sqrt{-\lambda}x} + C_2 e^{-\sqrt{-\lambda}x}.$$

For the boundary conditions (20.8) to be satisfied, it is necessary that

$$C_1 + C_2 = 0$$

and

$$C_1 e^{\sqrt{-\lambda}l} + C_2 e^{-\sqrt{-\lambda}l} = 0.$$

Consequently,

$$C_1 e^{\sqrt{-\lambda}l} = C_1 e^{-\sqrt{-\lambda}l}.$$

This last equation can be satisfied only if $C_1 = 0$ (and hence $C_2 = 0$). Then we obtain only a trivial solution of equation (20.7).

Case II ($\lambda = 0$). Then, the general solution of equation (20.7) is of the form

$$X(x) = C_1 + C_2 x.$$

For $X(0)$ to be equal to zero, it would be necessary for C_1 to be zero. But then, the condition $X(l) = 0$ takes the form $C_2 l = 0$. This would make it necessary that $C_2 = 0$. Thus, just as in the preceding case, we come to the conclusion

that only the trivial solution of equation (20.7) can satisfy both boundary conditions (20.8).

Case III ($\lambda > 0$). Here, the general solution of equation (20.7) takes the form

$$X(x) = C_1 \cos\sqrt{\lambda}x + C_2 \sin\sqrt{\lambda}x.$$

To satisfy the boundary condition $X(0) = 0$, we must have

$$C_1 = 0.$$

Then, the condition $X(l) = 0$ takes the form

$$C_2 \sin\sqrt{\lambda}l = 0 \quad \text{or} \quad \sin\sqrt{\lambda}l = 0,$$

so that, if C_2 were zero, we would again arrive at the trivial solution. The equation

$$\sin\sqrt{\lambda}l = 0$$

is satisfied if and only if

$$\sqrt{\lambda}l = k\pi, \quad \text{or} \quad \lambda = \frac{k^2\pi^2}{l^2},$$

where k is any integer (positive, negative, or zero). Since we assume that $\lambda > 0$, k cannot be zero. For negative values of k, the quantity λ takes the same values as for the corresponding positive values of k. Therefore, all values of λ for which equation (20.7) has a nontrivial solution satisfying the boundary conditions (20.8) are given by the formula

$$\lambda_k = \frac{k^2\pi^2}{l^2}, \text{where } k = 1, 2, \ldots \tag{20.9}$$

The problem of finding nontrivial solutions of equation (20.7) that satisfy the boundary conditions (20.8) is a special case of the problem known as the eigenvalue problem or the Sturm-Liouville problem (named after the two mathematicians who investigated it). Those values of λ at which our problem has nontrivial solutions are known as eigenvalues and the nontrivial solutions of the problem are called eigenfunctions corresponding to the given eigenvalues. In the present case, the eigenfunction

$$X_k(x) = C_k \sin\frac{k\pi}{l}x$$

corresponds to the eigenvalue $\frac{k^2\pi^2}{l^2}$.

Since equation (20.7) is homogeneous, the eigenfunctions are determined up to a constant factor C_k. If we choose this factor appropriately, we may subject the eigenfunction $X_k(x)$ to some additional condition, or, as we say, we may 'normalise' the eigenfunction.

This means that we should like to have

$$\int_0^l X_k^2(x)\,dx = 1.$$

For this, it is necessary that

$$C_k = \sqrt{\frac{2}{l}}.$$

Therefore we take (for the remainder of this section)

$$X_k(x) = \sqrt{\frac{2}{l}} \sin \frac{k\pi}{l} x.$$

Let us now turn to the solution of the mixed problem given at the beginning of the section. If we replace λ in equation (20.6) with its value λ_k, as given by formula (20.9), we obtain

$$T'' + \frac{k^2\pi^2}{l^2} T = 0.$$

We then have

$$T_k(t) = A_k \cos \frac{k\pi}{l} t + B_k \sin \frac{k\pi}{l} t,$$

where A_k and B_k are arbitrary constants.

All the functions

$$\begin{aligned} u_k(t, x) &= X_k(x)\, T_k(t) \\ &= \sqrt{\frac{2}{l}} \sin \frac{k\pi}{l} x \left(A_k \cos \frac{k\pi}{l} t + B_k \sin \frac{k\pi}{l} t \right) \end{aligned}$$

satisfy equation (20.1) and the boundary conditions (20.3) for arbitrary values of A_k and B_k. Let us seek values to assign to these constants such that the infinite series

$$u(t, x) = \sum_{k=1}^{\infty} X_k(x) \left[A_k \cos \frac{k\pi}{l} t + B_k \sin \frac{k\pi}{l} t \right] \qquad (20.10)$$

will satisfy equation (20.1), the boundary conditions (20.3), and the initial conditions (20.2). Let us start with the initial conditions. In the first place, it is necessary that

$$u(0,\ x)=\sum_{k=1}^{\infty} A_k X_k(x)=\varphi_0(x). \tag{20.11}$$

Furthermore, if the series can be differentiated termwise, we must have

$$u_t'(0,\ x)=\sum_{k=1}^{\infty} \frac{k\pi}{l} B_k X_k(x)=\varphi_1(x). \tag{20.12}$$

Let us assume that the functions $\varphi_0(x)$ and $\varphi_1(x)$ can be expanded in a series of terms of the form $\sin\frac{k\pi}{l}x$ on the interval $[0,\ l]$ such that the series formed by the absolute values of their terms converge uniformly.

We know from the theory of Fourier series that this is always possible if $\varphi_0(x)$ and $\varphi_1(x)$ and their first derivatives are continuous and if the values of these functions vanish at the end points of the interval $[0,\ l]$. Let us assume that these conditions are satisfied. Then, the series (20.10) converges absolutely and uniformly for $0\leqslant x\leqslant l$ and for arbitrary values of t since the functions $\sin\frac{k\pi}{l}t$ and $\cos\frac{k\pi}{l}t$ never exceed 1 in absolute value. From this, it follows that the function $u(t,\ x)$ defined by the series (20.10) is continuous and satisfies the first initial condition (20.2) and the boundary conditions (20.3). However, we should not conclude from this that this function satisfies the second initial condition (20.2) and equation (20.1). Such a conclusion might be made if the series (20.10) could be twice differentiated termwise with respect to x and twice with respect to t. As we know, termwise differentiation is lawful when the series obtained as a result of it converges uniformly in C_T. We know that this last condition will be satisfied for every T (including $T=\infty$) if (1) the first four derivatives of φ_0 exist and are continuous on the entire interval $[0,\ l]$, (2) φ_0 and its first two derivatives vanish at the end points of that interval, (3) the first three derivatives of the function φ_1 exist and are continuous on $[0,\ l]$, and (4) φ_1 and its first derivative vanish at the end

points of that interval*. In this case**

$$A_k = O\left(\frac{1}{k^4}\right) \text{ and } B_k = O\left(\frac{1}{k^4}\right)$$

2. In actual application of the Fourier method, one does not usually bother with the question of whether the series (20.10) can be twice differentiated termwise with respect to x and t. Instead, one merely checks to see that the functions φ_0 and φ_1 and their first derivatives are continuous and that these functions themselves vanish at the end points of the interval $[0, l]$. As we have seen, this ensures uniform and absolute convergence of the series (20.10) throughout the rectangle C_T. If for given values of φ_0 and φ_1 in the rectangle C_T, there exists a solution $u(t, x)$ of the problem in question that is continuous and has continuous derivatives of the first two orders, the sequence of partial sums $S_k(t, x)$ of the series (20.10) converges to it uniformly in C_T. This is true because, as we know from the theory of trigonometric series, the Fourier series corresponding to any function with integrable square converges to it in mean. Therefore, it follows from the way in which the series (20.10) was constructed that

$$\int_0^l [S'_{k_x}(0, x) - \varphi'_0(x)]^2\,dx \longrightarrow 0 \text{ and } \int_0^l [S'_{k_t}(0, x) - \varphi_1(x)]^2 dx \to 0$$

as $k \longrightarrow \infty$. It follows from this on the basis of Remark 3 to Section 19 that

$$S_k(t, x) \longrightarrow u(t, x)$$

uniformly in $\overline{C}_T$.

3. The so-called natural oscillations of a string fastened at both ends are described by each of the following three functions

* These restrictions on ϕ_0 and ϕ_1 can be weakened (see Section 23).

**By $O[\phi(n)]$ we denote a function $\psi(n)$ such that the ratio $\psi(n)/\phi(n)$ remains bounded as $n \to \infty$. These estimates are easily obtained if we transform the coefficients A_k and B_k by integration by parts.

$$u_k(t, x) = X_k(x)\,T_k(t)$$

$$= \sqrt{\frac{2}{l}} \sin\frac{k\pi}{l}x \left(A_k \cos\frac{k\pi}{l}t + B_k \sin\frac{k\pi}{l}t\right)$$

$$= D_k \sin\frac{k\pi}{l}x \sin\frac{k\pi}{l}(t + t_k) \quad (k = 1, 2, 3, \ldots)$$

When a string vibrates corresponding to $k = 1$, it emits the fundamental (lowest) tone. In the case of vibrations corresponding to higher values of k, it emits higher tones, known as overtones. If the string vibrates according to the law

$$u(t, x) = \sum_{k=1}^{n} D_k \sin\frac{k\pi}{l}x \sin\frac{k\pi}{l}(t + t_k),$$

it simultaneously emits sounds of different frequencies corresponding to the different terms in this sum.

21. THE GENERAL FOURIER METHOD (PRELIMINARY CONSIDERATIONS)

The Fourier method (or method of separation of variables as it is called) for solving a mixed boundary-value problem is applicable only to a certain special class of linear second-order equations, although the problem can be solved for a much wider class of equations.

In the present section, we shall explain the method without giving a rigorous justification of the results obtained. The justification will be given in subsequent sections. The first rigorous justification of the Fourier method was given by V.A. Steklov*.

Consider a hyperbolic equation of the form

$$A(t)\frac{\partial^2 u}{\partial t^2} + C(x)\frac{\partial^2 u}{\partial x^2} + D(t)\frac{\partial u}{\partial t} + E(x)\frac{\partial u}{\partial x} + [F_1(t) + F_2(x)]\,u = 0, \tag{21.1}$$

where the coefficients A, C, D, E, F_1, and F_2 are sufficiently smooth functions and where $A(t) > a_0 > 0$, and $C(x) < c_0 < 0$ (here, a_0 and c_0 are constants). The assumptions (1) that certain of these coefficients depend only on t, (2) that the

* Steklov's results are expounded in his book *Basic problems in mathematical physics*, Petrograd (1922).

others depend only on x, and (3) that the coefficient of $\frac{\partial^2 u}{\partial x \partial t}$ is equal to zero determine the class of hyperbolic equations for which the mixed boundary-value problem can be solved by the Fourier method.

Let us seek a twice continuously differentiable solution of equation (21.1) satisfying the initial conditions

$$u(0, x) = \varphi_0(x), \quad u_t'(0, x) = \varphi_1(x) \tag{21.2}$$

and the boundary conditions

$$\left.\begin{aligned} A_0 u(t, 0) + B_0 u_x'(t, 0) = 0, \\ A_1 u(t, l) + B_1 u_x'(t, l) = 0, \end{aligned}\right\} \tag{21.3}$$

where the constants A_0, B_0, A_1, and B_1 are such that $A_0^2 + B_0^2 \neq 0$ and $A_1^2 + B_1^2 \neq 0$.

As in the example of Section 20, let us first seek nontrivial solutions of equation (21.1) of the form

$$u(t, x) = T(t)\, X(x), \tag{21.4}$$

Here, we require that these solutions satisfy the boundary conditions (21.3). At the moment, we do not concern ourselves with satisfying the initial conditions.

If such a solution exists, when we substitute it into (21.1), we obtain an equation that the functions $X(x)$ and $T(t)$ must satisfy:

$$A(t)\, T''X + C(x)\, TX'' + D(t)\, T'X + E(x)\, TX' + [F_1(t) + F_2(x)]\, TX = 0.$$

Since the function $X(x)$ is not identically equal to zero, a point x_0 exists such that $X(x_0) \neq 0$. The equation

$$A(t)\, T'' + D(t)\, T' + F_1(t)\, T = -\frac{C(x_0)\, X''(x_0) + E(x_0)\, X'(x_0) + F_2(x_0)\, X(x_0)}{X(x_0)}\, T = \lambda_1 T,$$

where λ_1 is some constant, must be satisfied for all values of t. For the same reason, the function $X(x)$ must satisfy the equation

$$C(x)\, X'' + E(x)\, X' + F_2(x)\, X = \lambda_2 X,$$

where λ_2 is a constant, for all values of x. Since

$$A(t)\frac{T''}{T}+D(t)\frac{T'}{T}+F_1(t)$$
$$=-C(x)\frac{X''}{X}-E(x)\frac{X'}{X}-F_2(x) \qquad (21.5)$$

for all values of x and t such that $X(x)\neq 0$ and $T(t)\neq 0$, it follows that $\lambda_1=-\lambda_2=-\lambda$, and we obtain the following equations for the functions $X(x)$ and $T(t)$:

$$A(t)T''+D(t)T'+F_1(t)T+\lambda T=0, \qquad (21.6)$$
$$C(x)X''+E(x)X'+F_2(x)X-\lambda X=0. \qquad (21.7)$$

Since $T(t)\not\equiv 0$ for the function (21.4) to satisfy the boundary conditions (21.3), it is necessary that

$$\left.\begin{aligned} A_0X(0)+B_0X'(0)&=0,\\ A_1X(l)+B_1X'(l)&=0.\end{aligned}\right\} \qquad (21.8)$$

The problem of finding nontrivial solutions of equation (21.7) satisfying the conditions (21.8) is called an eigenvalue problem. This problem does not have a nontrivial solution (that is, one not identically equal to zero) for all values of λ. Those values of λ at which a nontrivial solution exists are called eigenvalues of the problem, and the nontrivial solution itself is called the eigenfunction corresponding to the eigenvalue in question. The set of all eigenvalues is called the spectrum of the given problem.

In the following section, we shall show that the eigenvalues of our problem constitute an infinite sequence

$$\lambda_1, \lambda_2, \ldots, \lambda_k, \ldots$$

To each eigenvalue λ_k there corresponds an eigenfunction $X_k(x)$ which, because of the homogeneity of equation (21.7) and conditions (21.8), is defined up to an arbitrary numerical factor. Let us choose this factor so that

$$\int_0^l \rho(x)X_k^2(x)\,dx=1, \qquad (21.9)$$

where $\rho(x)>0$ is some particular function corresponding to the given equation. This function will be determined in the following section.

It will be shown later that the eigenfunctions corresponding

to different eigenvalues are 'orthogonal with weight ρ'; that is, they satisfy the equations

$$\int_0^l \rho(x) X_k(x) X_l(x)\, dx = 0 \quad \text{for} \quad k \neq l. \tag{21.10}$$

Let us find the solution of equation (21.6) corresponding to each eigenvalue λ_k. The general solution of equation (21.6) for $\lambda = \lambda_k$ (which we denote by $T_k(t)$) is an arbitrary linear combination of any two linearly independent particular solutions $T_k^*(t)$ and $T_k^{**}(t)$:

$$T_k(t) = C_1 T_k^*(t) + C_2 T_k^{**}(t).$$

Let us choose T_k^* and T_k^{**} that will satisfy the following initial conditions at $t = 0$:

$$\left.\begin{aligned} T_k^*(0) = 1; \quad T_k^{*\prime}(0) = 0; \\ T_k^{**}(0) = 0; \quad T_k^{**\prime}(0) = 1, \end{aligned}\right\} \tag{21.11}$$

and let us set

$$u_k(t, x) = T_k(t) \cdot X_k(x).$$

For arbitrary k, the functions $u_k(t, x)$ satisfy equation (21.1) and the boundary conditions (21.3).

To satisfy the initial conditions (21.2), we set up the series

$$u(t, x) = \sum_{k=1}^{\infty} X_k(x) \left[A_k T_k^*(t) + B_k T_k^{**}(t)\right]. \tag{21.12}$$

If this series converges uniformly and if the series obtained by twice differentiating it termwise with respect to t and to x converge uniformly, its sum will clearly satisfy equation (21.1) and the boundary conditions (21.3). For the initial conditions (21.2) to be satisfied, it is necessary that

$$u(0, x) = \sum_{k=1}^{\infty} A_k X_k(x) = \varphi_0(x), \tag{21.13}$$

$$u_t'(0, x) = \sum_{k=1}^{\infty} B_k X_k(x) = \varphi_1(x). \tag{21.14}$$

Assuming that the series (21.13) and (21.14) converge uniformly, we can determine the coefficients A_m and B_m by multiplying both sides of these two equations by $\rho X_m(x)$ and integrating from 0 to l with respect to x. Because of (21.9) and (21.10), we obtain

$$A_m = \int_0^l \rho(x) \varphi_0(x) X_m(x)\, dx,$$

$$B_m = \int_0^l \rho(x) \varphi_1(x) X_m(x)\, dx.$$

When we substitute these values of the coefficients into the series (21.12), we clearly obtain the solution of our problem if both the series (21.12) and the series that are obtained from it by differentiating twice termwise with respect to x and t converge uniformly.

Remark: We have shown the general procedure for applying the Fourier method to solving the mixed problem for equation (21.1). This procedure can also be applied in the case of hyperbolic equations of a special form (see Section 25) with several space variables.

22. GENERAL PROPERTIES OF EIGENFUNCTIONS AND EIGENVALUES

1. In studying the properties of eigenfunctions and eigenvalues, let us show first of all that equation (21.7)

$$C(x) X''(x) + E(x) X'(x) + F_2(x) X(x) - \lambda X(x) = 0$$

of the preceding section can be reduced to the form

$$[p(x) X'(x)]' - q(x) X(x) + \lambda \rho(x) X(x) = 0, \qquad (22.1)$$

by multiplying it by a suitably chosen function of x.

In all that follows, we shall assume that $C(x) < c_0 < 0$, where c_0 is a constant. If we then multiply (21.7) by $\rho(x)$, we obtain

$$\rho C X'' + \rho E X' + \rho F_2 X - \lambda \rho X = 0.$$

For the first two terms to be representable in the form

$$[p(x) X']',$$

it is necessary that

$$(\rho C)' = \rho E.$$

Determining $\rho(x)$ from this differential equation, we obtain

$$\rho(x) = e^{\int\limits_{x_0}^{x} \frac{E-C'}{C} dx} > 0.$$

Here we have taken a particular solution of the differential equation for $\rho(x)$. We define

$$\rho C = -p \text{ and } \rho F_2 = q.$$

We can now write our equation in the form (22.1). It follows from the assumptions made that $p(x) > p_0$ and $\rho(x) > \rho_0$, where p_0 and ρ_0 are positive constants.

We shall assume that $p(x)$, $p'(x)$, $q(x)$ and $\rho(x)$ are continuous for $0 \leqslant x \leqslant l$.

2. Thus, we have an eigenvalue problem to investigate, that of finding a nontrivial solution of equation (22.1) satisfying the conditions

$$\left.\begin{aligned} A_0 X(0) + B_0 X'(0) &= 0, \\ A_1 X(l) + B_1 X'(l) &= 0, \end{aligned}\right\} \qquad (22.2)$$

where $A_0^2 + B_0^2 \neq 0$ and $A_1^2 + B_1^2 \neq 0$.

Theorem 1. If $X_1(x)$ and $X_2(x)$ are eigenfunctions corresponding to the same eigenvalue λ, then $X_1(x) = cX_2(x)$, where c is a constant.

Proof: Since, by assumption, $X_1(x)$ and $X_2(x)$ satisfy the conditions

$$A_0 X_1(0) + B_0 X_1'(0) = 0,$$
$$A_0 X_2(0) + B_0 X_2'(0) = 0$$

and $A_0^2 + B_0^2 \neq 0$, the Wronskian determinant

$$\begin{vmatrix} X_1 & X_2 \\ X_1' & X_2' \end{vmatrix}$$

of the solutions X_1 and X_2 of equation (22.1) must vanish

at the point $x=0$, which means that the functions $X_1(x)$ and $X_2(x)$ are linearly dependent.

In what follows, we shall assume that the eigenfunctions are normalised with weight ρ, i.e. they are chosen so that

$$\int_0^l \rho(x)[X(x)]^2\,dx = 1. \tag{22.3}$$

Such a function $X(x)$ can be obtained by multiplying an arbitrary eigenfunction $\tilde{X}(x)$ by the number

$$\frac{1}{\sqrt{\int_0^l \rho(x)[\tilde{X}(x)]^2\,dx}}.$$

Obviously, with the given eigenvalue, the normalised eigenfunction is determined except for sign.

Theorem 2. The eigenfunctions corresponding to different eigenvalues are orthogonal with weight $\rho(x)$; that is, if $\lambda_1 \neq \lambda_2$ and if $X_i(x)$ is the eigenfunction corresponding to the eigenvalue λ_i (for $i=1, 2$), then

$$\int_0^l \rho(x)X_1(x)X_2(x)\,dx = 0. \tag{22.4}$$

Proof: Consider the identities

$$(pX_1')' - qX_1 + \lambda_1\rho X_1 = 0, \quad (pX_2')' - qX_2 + \lambda_2\rho X_2 = 0.$$

If we multiply the first of these by X_2 and the second by X_1 and then subtract the second from the first, we obtain the identity

$$[pX_1']'X_2 - [pX_2']'X_1 + (\lambda_1 - \lambda_2)\rho X_1X_2 = 0.$$

If we integrate this identity from 0 to l (and then integrate the first two terms by parts), we obtain

$$\int_0^l (\lambda_2 - \lambda_1)\rho X_1X_2\,dx$$

$$= pX_1'X_2\Big|_0^l - pX_2'X_1\Big|_0^l - \int_0^l pX'X'\,dx + \int_0^l pX_2'X_1'\,dx.$$

The right-hand side of this equation is equal to zero since the last two terms cancel each other out and since

$$p(l)\,[X_1'(l)\,X_2(l) - X_2'(l)\,X_1(l)] = 0$$

and

$$p(0)\,[X_1'(0)\,X_2(0) - X_2'(0)\,X_1(0)] = 0$$

because of conditions (22.2). Therefore,

$$\int_0^l \rho X_1 X_2\, dx = 0,$$

since $\lambda_1 \neq \lambda_2$.

3. To simplify the exposition that follows, we shall confine ourselves to consideration of the boundary conditions

$$\begin{aligned} X(0) &= 0, \\ X(l) &= 0. \end{aligned} \tag{22.5}$$

In this section, the problem of finding the eigenvalues will be reduced to the problem of finding a conditional extremum (specifically, a minimum) of some functional. This functional will be chosen in such a way that equation (22.1) will be the Euler-Lagrange equation for it*.

Consider the two quadratic functionals of the function $X(x)$

$$D(X) = \int_0^l (pX'^2 + qX^2)\, dx,$$

$$H(X) = \int_0^l \rho X^2\, dx.$$

The functionals

$$D(X_1,\ X_2) = \int_0^l (pX_1'X_2' + qX_1X_2)\, dx,$$

* See LAVRENT'EV, M.A. and LYUSTERNIK, L.A., *Course in the calculus of variations*, 2nd. ed. Gostekhizdat (1950).

$$H(X_1, X_2) = \int_0^l \rho X_1 X_2 \, dx$$

are called bilinear functionals, corresponding to the given quadratic ones. For these we have the following theorem.

Theorem 3. If λ is an eigenvalue of the eigenvalue problem that we are considering and if X_λ is the normalised eigenfunction corresponding to it, then for any continuously differentiable function $f(x)$ satisfying equations (22.5),

$$D(X_\lambda, f) = \lambda \int_0^l \rho X_\lambda f \, dx = \lambda H(X_\lambda, f).$$

Proof: Integrating by parts and using conditions (22.5) for the function f and equation (22.1), we obtain

$$\begin{aligned} D(X_\lambda, f) &= \int_0^l (pX_\lambda' f' + qX_\lambda f) \, dx \\ &= pX_\lambda' f \Big|_0^l - \int_0^l [(pX_\lambda')' - qX_\lambda] f \, dx = \lambda \int_0^l \rho X_\lambda f \, dx \\ &= \lambda H(X_\lambda, f). \end{aligned}$$

Corollary. Suppose that $X_i(x)$ is the eigenfunction corresponding to the eigenvalue λ_i then,

$$D(X_i) = \lambda_i, \quad D(X_i, X_j) = 0, \quad \text{if} \quad i \neq j.$$

Theorem 4. For $X \not\equiv 0$ the ratio $\frac{D(X)}{H(X)}$ is bounded below and, consequently, we have an exact lower bound.

This is true because

$$\begin{aligned} D(X) &= \int_0^l (pX'^2 + qX^2) \, dx \geqslant \int_0^l qX^2 \, dx \\ &= \int_0^l \frac{q}{\rho} \rho X^2 \, dx \geqslant \min_{0 \leqslant x \leqslant l} \frac{q(x)}{\rho(x)} H(X). \end{aligned}$$

If we consider functions that satisfy the conditions $H(X) = 1$, the values of the functional $D(X)$ itself will be

bounded below for these functions. And since every eigenvalue $\lambda=D(X_\lambda)$ if $H(X_\lambda)=1$, we obtain the important consequence of this that the eigenvalues of our problem are bounded below.

Let us consider the problem of finding the minimum of the functional $D(X)$ under the condition that $H(X)=1$. We take as the class of admissible functions the set of functions $X(x)$ that are twice continuously differentiable on the closed interval $0\leqslant x\leqslant l$ and satisfy conditions (22.5). We assume that this minimum is attained in the class of admissible functions*. Then, as we know from the calculus of variations, the function representing this minimum must satisfy for some λ the Euler-Lagrange equation for the functional

$$D(X)-\lambda H(X)=\int_0^l [pX'^2+qX^2-\lambda\rho X^2]\,dx$$
$$=\int_0^l F(x,\ X,\ X')\,dx,$$

that is, an equation of the form

$$\frac{\partial F}{\partial X}-\frac{d}{dx}\frac{\partial F}{\partial X'}=0,$$

which in our case coincides with equation (22.1). The boundary conditions of the eigenvalue problem and the variational problem in question also coincide. Therefore, the function $X_1(x)$ representing the extremum $D(X)$ under the condition that $H(X)=1$ is an eigenfunction of the original eigenvalue problem. Since, according to Theorem 3, $D(X_\lambda)=\lambda$, the eigenvalue corresponding to $X_1(x)$ must obviously be the smallest one. We denote it by λ_1.

* The proof of the existence of a solution to this problem and of all the other variational problems of which we shall speak in the present chapter is given in the supplement by GEL'FAND, I.M. and SUKHOMLINOV, G.A., to the book by LAVRENT'EV, M.A. and LYUSTERNIK, L.A., *Fundamentals of calculus of variations*, I, Chapter II (1935).

It is shown in the calculus of variations that if we required only the existence of continuous first derivatives of the admissible functions, the variational problem that we are considering would still have a solution. Its solution would necessarily have continuous second-order derivatives and therefore would coincide with the solution of the same problem in the case of twice continuously differentiable functions.

Let us show that the function $X(x)$ representing the minimum of the functional $D(X)$ in the class of admissible functions that satisfy all the original conditions and in addition the supplementary condition

$$\int_0^l \rho X_1 X\,dx = 0,$$

is the eigenfunction corresponding to the second smallest eigenvalue.

The function representing the minimum in question must satisfy the Euler-Lagrange equation for the functional

$$D(X) - \lambda H(X) - \mu \int_0^l \rho X_1 X\,dx,$$

which, in the present case, may be written in the form

$$(pX')' - qX + \lambda\rho X + \frac{1}{2}\mu\rho X_1 = 0; \tag{22.6}$$

where λ and μ are constants.

Let us show that $\mu = 0$. To show this, we rewrite equation (22.1), replacing λ with λ_1 and X with X_1:

$$(pX_1')' - qX_1 + \lambda_1\rho X_1 = 0. \tag{22.7}$$

Let us multiply (22.6) by X_1 and (22.7) by X and then subtract one of the resulting equations from the other and integrate from 0 to l. Repeating the integration by parts that was performed in the proof of the orthogonality and using the fact that

$$\int_0^l \rho X_1 X\,dx = 0,$$

we obtain

$$\mu \int_0^l \rho X_1^2\,dx = 0,$$

from which if follows that

$$\mu = 0.$$

Consequently, equation (22.6) takes the form

$$(pX')' - qX + \lambda\rho X = 0,$$

and X is thus an eigenfunction. Let us denote it by $X_2(x)$. Let us show that the eigenvalue λ_2 corresponding to this function is the closest eigenvalue to λ_1. Obviously, $\lambda_2 \geqslant \lambda_1$, since increase in the number of conditions on the admissible functions can only cause the minimum $D(X)$ to increase. The value of λ_2 cannot be equal to λ_1 since then $X_2(x)$ would, from Theorem 1, be equal to $\pm X_1(x)$ which contradicts the condition that $\int\limits_0^l X_1 X\,dx = 0$. Consequently, $\lambda_2 > \lambda_1$.

Let us show that there are no eigenvalues between λ_2 and λ_1. If there were three eigenvalues $\lambda_2 > \tilde{\lambda} > \lambda_1$ corresponding to the eigenfunctions X_2, $\tilde{X}$, and X_1, then, as we can easily see, it would not be the function X_2 but the function $\tilde{X}$ that would, according to the corollary of Theorem 3, be the solution of the variational problem that we just examined.

In an analogous way, it can be shown that the function $X_n(x)$ representing the minimum of $D(X)$ in the class of twice continuously differentiable functions satisfying the relationships (22.5) and the conditions

$$H(X) = 1, \quad \int\limits_0^l \rho X_j X\,dx = 0 \quad (j = 1, 2, \ldots, n-1),$$

where $X_j(x)$ is the jth eigenfunction, is the eigenfunction corresponding to the nth eigenvalue.

Thus, we have given the method for finding successively the eigenvalues and eigenfunctions. Since we shall show later that $\lambda_n \to \infty$ as $n \to \infty$, it follows that all the eigenvalues and eigenfunctions can be found in this way.

4. A method exists for finding the nth eigenvalue and the nth eigenfunction without first finding the preceding eigenfunctions. This method is indicated by the following theorem by Courant:

Theorem. Suppose that $\varphi_1(x), \varphi_2(x), \ldots, \varphi_{n-1}(x)$ form an arbitrary system of continuous functions on the interval $[0, l]$. Let us denote by $\lambda(\varphi_1, \ldots, \varphi_{n-1})$ the minimum of

the functional $D(X)$ in the class of twice continuously differentiable functions that vanish at the end points of the interval and that satisfy the following additional conditions:

$$H(X)=1, \tag{22.8}$$

$$\int_0^l \rho\varphi_i X\,dx=0 \quad (i=1,\ 2,\ \ldots,\ n-1). \tag{22.9}$$

Then, the n th eigenvalue λ_n for the above eigenvalue problem is equal to the least upper bound of the values $\lambda(\varphi_1, \varphi_2, \ldots, \varphi_{n-1})$ corresponding to all possible choices of the functions $\varphi_1, \ldots, \varphi_{n-1}$

Proof: From what was said above,

$$\lambda(X_1, \ldots, X_{n-1})=\lambda_n,$$

so that it will be sufficient to show that, for an arbitrary choice of $\varphi_1, \ldots, \varphi_{n-1}$

$$\lambda(\varphi_1, \ldots, \varphi_{n-1})\leqslant\lambda_n.$$

Let us show that, for an arbitrary system $\varphi_1, \ldots, \varphi_{n-1}$ and admissible function $\breve{X}(x)$ can be exhibited which satisfies conditions (22.5), (22.8) and (22.9) and that

$$D(\breve{X})\leqslant\lambda_n.$$

It will follow from this that

$$\lambda(\varphi_1, \ldots, \varphi_{n-1})\leqslant\lambda_n$$

which will prove the theorem.

Let us seek a function $\breve{X}(x)$ in the form

$$\breve{X}(x)=\sum_{k=1}^{n} c_k X_k(x).$$

Clearly, such a function will vanish at $x=0$ and $x=l$ for arbitrary values of c_k and it will have continuous first and second derivatives. Let us choose the coefficients c_k in such a way that conditions (22.8) and (22.9) will be satisfied. When we substitute $\breve{X}$ into (22.8) and use the fact that $H(X_i, X_k)=0$ for $i\neq k$ (the property of orthogonality of eigenfunctions), we obtain

$$H(\tilde{X}) = \int_0^l \rho \tilde{X}^2 \, dx = \sum_{k=1}^n c_k^2 = 1. \tag{22.10}$$

Conditions (22.9) yield the system of equations

$$\int_0^l \rho \varphi_i \tilde{X} \, dx = \sum_{k=1}^n c_k \int_0^l \rho \varphi_i X_k \, dx = 0 \qquad (i = 1, 2, \ldots, n-1).$$

This is a system of $n-1$ linear equations with n unknowns c_k. It always has nontrivial solutions. When we normalise one of these solutions by use of (22.10), we choose the function $\tilde{X}(x)$. We find $D(\tilde{X})$:

$$\begin{aligned} D(\tilde{X}) &= \int_0^l \left[p \left(\sum_{k=1}^n c_k X_k' \right)^2 + q \left(\sum_{k=1}^n c_k X_k \right)^2 \right] dx \\ &= \int_0^l \left(p \sum_{k=1}^n \sum_{l=1}^n c_k c_l X_k' X_l' + q \sum_{k=1}^n \sum_{l=1}^n c_k c_l X_k X_l \right) dx \\ &= \sum_{k=1}^n c_k^2 D(X_k) + \sum_{k \neq l} c_k c_l D(X_k, X_l) \\ &= \sum_{k=1}^n c_k^2 D(X_k) = \sum_{k=1}^n c_k^2 \lambda_k \leqslant \lambda_n \sum_{k=1}^n c_k^2 \quad = \lambda_n, \end{aligned}$$

which completes the proof†.

Remark: Instead of finding the minimum of the functional $D(X)$ under the conditions (22.8) and (22.9), we can seek the minimum of the ratio $\frac{D(X)}{H(X)}$ under conditions (22.9). The minimum value will be the same in both cases. However, in the second case, the function representing the extreme value will be determined up to a constant factor.

5. For later investigation of eigenvalues and eigenfunctions, let us show how the eigenvalues change when the coefficients in equation (22.1) and the interval on which solutions are being considered change,

(a) When the coefficients $p(x)$ and $q(x)$ are both increased or decreased, the eigenvalues also increase or decrease respectively. More precisely, if we have two equations

† Since $\lambda_1 < \lambda_2 < \ldots < \lambda_n$.

$$(pX')' - qX + \lambda\rho X = 0,$$

$$(p^*X')' - q^*X + \lambda\rho X = 0,$$

where

$$p(x) \leqslant p^*(x), \quad q(x) \leqslant q^*(x),$$

then $\lambda_n \leqslant \lambda_n^*$, where λ_n and λ_n^* are respectively the n th eigenvalues of the first and second equations above.

The proof follows directly from the fact that

$$D(X) = \int_0^l (pX'^2 + qX^2)\,dx \leqslant \int_0^l (p^*X'^2 + q^*X^2)\,dx = D^*(X).$$

Therefore,

$$\lambda(\varphi_1, \ldots, \varphi_{n-1}) \leqslant \lambda^*(\varphi_1, \ldots, \varphi_{n-1}),$$

since the class of admissible functions $X(x)$ has not changed and, consequently, $\lambda_n \leqslant \lambda_n^*$.

(b) When the coefficient $\rho(x)$ is increased or decreased, the eigenvalues decrease or increase respectively (that is, they change in the opposite way).

Suppose that $\rho(x) \leqslant \rho^*(x)$, and that the remaining coefficients in the equation are unchanged. Then, for every function $X(x)$,

$$D(X) = D^*(X), \text{ and } H(X) \leqslant H^*(X).$$

Therefore,

$$\frac{D(X)}{H(X)} \geqslant \frac{D^*(X)}{H^*(X)}. \tag{22.11}$$

Every function $X(x)$ satisfying conditions (22.9) for certain $\varphi_i(x)$ will satisfy the analogous conditions

$$\int_0^l \rho^*(x)\varphi_i^*(x)\,X(x)\,dx = 0,$$

if we set

$$\varphi_i^*(x) = \frac{\rho(x)}{\rho^*(x)}\varphi_i(x).$$

From this and from inequality (22.11), we conclude that

$$\lambda(\varphi_1, \ldots, \varphi_{n-1}) \geqslant \lambda^*(\overset{*}{\varphi}_1, \ldots, \overset{*}{\varphi}_{n-1}).$$

Since $\rho^*(x) \geqslant \rho(x) > \rho_0 > 0$, the set of all $(\varphi_1(x), \ldots, \varphi_n(x))$ coincides with the set of all $(\overset{*}{\varphi}_1(x), \ldots, \overset{*}{\varphi}_n(x))$, and therefore, $\lambda_n \geqslant \overset{*}{\lambda}_n$.

(c) The eigenvalues do not decrease when the length of the interval $[0, l]$ is decreased. More precisely, if we replace the interval $[0, l]$ in the eigenvalue problem in question with an interval $[0, l^*]$, where $l^* < l$ and the eigenvalues of the new problem are denoted by λ^*, then $\lambda_n^* \geqslant \lambda_n$.

Proof: $\lambda^*(\varphi_1, \ldots, \varphi_{n-1})$, which plays the same role in the new problem that $\lambda(\varphi_1, \ldots, \varphi_{n-1})$ played in the old, will coincide with the minimum of the functional $D(X)$ defined for the interval $[0, l]$ if we impose on the class of admissible functions $X(x)$ not only the conditions (22.8) and (22.9) but also the requirement that $X(x) \equiv 0$ for $l^* \leqslant x \leqslant l$. However, when the conditions are made more stringent, the class of admissible functions becomes smaller and the minimum of the functional can only decrease†.

Consequently,

$$\lambda^*(\varphi_1, \ldots, \varphi_{n-1}) \geqslant \lambda(\varphi_1, \ldots, \varphi_{n-1}).$$

Therefore,

$$\overset{*}{\lambda}_n \geqslant \lambda_n.$$

Let us look at the specific example considered in Section 20. We can derive from it the familiar relationship between the length of a string and the frequency of its basic tone: the shorter the string, the higher will be the frequency of its natural vibrations (equal to $\frac{k\pi}{l}$) and the higher will be the note emitted by it

6. In just the same way as we investigated the eigenvalues for the problem (22.1) under the boundary conditions

† Requiring that the functions $X(x)$ which are continuous and have continuous first and second derivatives on the interval $[0, l]$ vanish for $l^* \leqslant x \leqslant l$, amounts to requiring that not only the function $X(x)$ but also its first two derivatives, vanish at $x = l^*$. However, one can show that this additional requirement does not change the minimum of $D(X)$ on the interval $[0, l^*]$.

$$X(0)=0, \quad X(l)=0, \tag{22.5}$$

we can investigate the eigenvalues for equation (22.1) with the boundary conditions

$$X'(0)=0, \quad X'(l)=0, \tag{22.12}$$

or with the boundary conditions

$$X'(0)-\sigma_0 X(0)=0, \quad X'(l)+\sigma_l X(l)=0, \tag{22.13}$$

where $\sigma_0 \geqslant 0$ and $\sigma_l \geqslant 0$, or with a boundary condition of one of the types shown here imposed on one end point of the interval $(0, l)$ and one of the other type at the other end.

The fundamental theorem making it possible to investigate the eigenvalues under the boundary conditions (22.13) is the following theorem, which is analogous to the theorem of subsection 4.

Theorem. Suppose that $\varphi_1(x), \varphi_2(x), \ldots, \varphi_{n-1}(x)$ is an arbitrary system of continuous functions on the interval $[0, l]$. Let us denote by $\lambda(\varphi_1, \ldots, \varphi_{n-1})$ the minimum of the functional

$$\int_0^l (pX'^2+qX^2)\,dx+\sigma_0 p(0) X^2(0)+\sigma_l p(l) X^2(l) \tag{22.14}$$

in the class of twice continuously differentiable functions satisfying the following conditions:

$$H(X)=1, \quad \int_0^l \rho\varphi_i X\,dx=0 \qquad (i=1, 2, \ldots, n-1). \tag{22.15}$$

Then, the n th eigenvalue λ_n for this eigenvalue problem is equal to the least upper bound of the values $\lambda(\varphi_1, \ldots, \varphi_{n-1})$ for all possible choices of continuous functions $\varphi_1, \ldots, \varphi_{n-1}$.

By use of this theorem, we may, just as in the case of fixed end points, investigate the dependence of the eigenvalues on $p(x)$, $q(x)$, $\rho(x)$, σ_0, σ_l, l.

If for the functions $\varphi_1, \ldots, \varphi_{n-1}$ we take the first $n-1$ eigenfunctions $X_1, \ldots, X_{n-1}$ of the problem in question, then the function representing the minimum of the functional (22.14) under conditions (22.15) will be the n th eigenfunction of this problem and the minimum of the function will be its

n th eigenvalue.

If $\sigma_0=0$ and $\sigma_l=0$, we arrive at the eigenvalue problem for equation (22.1) with the boundary conditions (22.12). In this case, the nth eigenfunction will yield the minimum of $D(X)$ in the class of twice continuously differentiable functions satisfying the same conditions

$$H(X)=1, \int_0^l \rho X_j X\,dx=0 \qquad (j=1, 2, \ldots, n-1),$$

where $X_1, \ldots, X_{n-1}$ are the first eigenfunctions of this problem, as in the case of fixed endpoints. Now, however, we require of the admissible functions that they satisfy some condition or other on the end points of the interval $(0, l)$. The function that satisfies this variational problem will automatically satisfy conditions (22.12). This is the 'free problem'. It corresponds to the vibrations of a free string, that is, one whose end points are not fixed. However, let us recall that when we say that the string is not fixed at the end points, this means only that these ends may move freely along a straight line perpendicular to the position of equilibrium of the string. It emphatically does not mean that these ends can move along the equilibrium position of the string.

If we do not require that the admissible functions be continuous at any interior point c of the interval $(0, l)$, the class of admissible functions will be made broader. Then, $\lambda(\varphi_1, \varphi_2, \ldots, \varphi_{n-1})$, and, consequently, λ_n can only be decreased as a result of this. The corresponding tone emitted by the string will be lower. This corresponds to a discontinuity in the string at some interior point c. Then, the ends of both portions of the string (each moving along a single straight line perpendicular to the stationary position of the stretched string) can move freely along that line. The corresponding eigenfunction X_n will have a discontinuity of the first kind at the point c. In this case,

$$X_n'(c+0)=0 \quad \text{and} \quad X_n'(c-0)=0.$$

It follows from what was said above that the tones emitted by the string will be lower in this case than the tones of the unbroken string.

7. Again, we shall confine ourselves to a consideration of boundary conditions of the type (22.5) since completely ana-

logous reasoning can be used in the other cases.

Let us make an estimate of λ_n, depending on n. We denote the maxima of the functions $p(x)$, $q(x)$ and $\rho(x)$ on the interval $[0, l]$ by p_{max}, q_{max}, and ρ_{max} respectively, and we denote their minima by p_{min}, q_{min}, and ρ_{min}. In addition to equation (22.1), let us look at the equations with constant coefficients

$$p_{max}X'' - q_{max}X + \lambda\rho_{min}X = 0, \tag{22.16}$$

$$p_{min}X'' - q_{min}X + \lambda\rho_{max}X = 0. \tag{22.17}$$

It follows from the results of subsection 5 that

$$\underline{\lambda}_n \leqslant \lambda_n \leqslant \overline{\lambda}_n, \tag{22.18}$$

where $\overline{\lambda}_n$ and $\underline{\lambda}_n$ are the nth eigenvalues of equations (22.16) and (22.17) respectively. However, equations (22.16) and (22.17) can be integrated in finite form and the values of $\overline{\lambda}_n$ and $\underline{\lambda}_n$ can be calculated precisely. If we solve (22.16), for example, and find a particular solution of this equation from the conditions $X(0) = X(l) = 0$, we obtain

$$\frac{\overline{\lambda}_n\rho_{min} - q_{max}}{p_{max}} = \frac{n^2\pi^2}{l^2}.$$

Therefore, $\overline{\lambda}_n = C_1n^2 + C_2$, where C_1 and C_2 are independent of n.

Analogously,

$$\underline{\lambda}_n = c_1n^2 + c_2.$$

When we substitute these values into (22.18), we obtain

$$c_1n^2 + c_2 \leqslant \lambda_n \leqslant C_1n^2 + C_2. \tag{22.19}$$

From this it follows in particular that the eigenvalues increase without bound as $n \to \infty$.

8. Let us now investigate the behaviour of the eigenfunctions with increase in n. To do this, we simplify equation (22.1) by making the substitution

$$s = \int_0^x \varphi(x)\,dx, \quad u = \frac{1}{\psi(x)} X. \tag{22.20}$$

We choose the functions $\varphi(x) > 0$ and $\psi(x) > 0$ in such a way that, after the substitution (22.20), equation (22.1) will become

$$u''(s) + \lambda u = R(s)\,u. \tag{22.21}$$

When we carry out the substitution (22.20) with arbitrary functions $\varphi(x)$ and $\psi(x)$, we shift from equation (22.1) to the equation

$$\frac{d^2u}{ds^2} + \frac{(\varphi\psi p)' + \varphi\psi' p}{\varphi^2\psi p}\frac{du}{ds} + \lambda\rho\frac{1}{\varphi^2 p}u = \frac{\psi q - (\psi' p)'}{\varphi^2\psi p}u.$$

Let us now choose functions $\varphi(x)$ and $\psi(x)$ in such a way that this equation will take the form (22.21). For this, we need to determine the functions $\varphi(x)$ and $\psi(x)$ from the system of equations

$$\frac{\rho}{\varphi^2 p} = 1, \quad (\varphi\psi p)' + \varphi\psi' p = 0.$$

When we solve this system, we obtain

$$\varphi = \sqrt{\frac{\rho}{p}}, \quad \psi = \frac{c}{\sqrt[4]{\rho p}},$$

where c is an arbitrary constant.

Therefore, by making the change

$$s = \int_0^x \sqrt{\frac{\rho}{p}}\,dx, \quad u = \sqrt[4]{\rho p}\,X \tag{22.22}$$

for example, we can obtain equation (22.21). Here, $R(s)$ is a continuous function if $\rho''(x)$ and $p''(x)$ are continuous since $\varphi^2\psi p \neq 0$.

We need to find a solution of equation (22.21) on the interval

$$0 < s < l_1, \text{where } l_1 = \int_0^l \sqrt{\frac{\rho}{p}}\,dx.$$

The boundary conditions for $u(s)$ remain the same as for $X(x)$, as can easily be seen:

$$u(0)=0, \quad u(l_1)=0.$$

If $X_n(x)$ is the eigenfunction of equation (22.1) corresponding to the eigenvalue λ_n, then the eigenfunction u_n of the equation (22.21) corresponds to this same eigenvalue.

If

$$\int_0^l \rho X_n^2 dx = 1,$$

then, as can easily be seen,

$$\int_0^{l_1} u_n^2(s) ds = 1. \tag{22.23}$$

Let us give asymptotic formulae for $u_n(s)$ as $n \to \infty$. Since we are concerned with the behaviour of $u_n(s)$ for large values of n, we may, on the basis of (22.19), confine ourselves to positive λ_n. Let us examine the nonhomogeneous equation for the function $z(s)$

$$z'' + \lambda z = R(s) u, \quad \lambda > 0, \tag{22.24}$$

where $u(s)$ is a solution of equation (22.21) for the same λ. The general solution of equation (22.24) is

$$z(s) = C_1 \cos \sqrt{\lambda} s + C_2 \sin \sqrt{\lambda} s + \frac{1}{\sqrt{\lambda}} \int_0^s R(\tau) u(\tau) \sin \sqrt{\lambda} (s-\tau) d\tau$$

If we set $C_1 = u(0)$ and $C_2 = \frac{u'(0)}{\sqrt{\lambda}}$, then $z(s)$ will, for $s=0$, satisfy the same initial conditions as does $u(s)$. Therefore, on the basis of the uniqueness theorem for the solution of the Cauchy problem for equation (22.24), $z(s)$ will be identically equal to $u(s)$, and we will obtain for $u(s)$ the integral equation

$$u(s) = u(0) \cos \sqrt{\lambda}\, s + \frac{u'(0)}{\sqrt{\lambda}} \sin \sqrt{\lambda}\, s + \frac{1}{\sqrt{\lambda}} \int_0^s R(\tau) u(\tau) \sin \sqrt{\lambda}\, (s-\tau) d\tau \tag{22.25}$$

Now, suppose that λ coincides with the nth eigenvalue and

that $v_n(s)$ is the solution to equation (22.21) at $\lambda = \lambda_n$ satisfying the initial conditions

$$v_n(0) = 0; \quad v_n'(0) = \sqrt{\lambda_n}.$$

This function $v_n(s)$ will satisfy the integral equation

$$v_n(s) = \sin\sqrt{\lambda_n}\, s + \frac{1}{\sqrt{\lambda_n}} \int_0^s R(\tau)\, v_n(\tau) \sin\sqrt{\lambda_n}(s-\tau)\, d\tau. \tag{22.26}$$

The function $v_n(s)$ differs from the normalised eigenfunction $u_n(s)$ only by sign and by the factor

$$u_n(s) = \frac{v_n(s)}{\sqrt{\int_0^{l_1} v_n^2(s)\, ds}} = N_n v_n.$$

We shall show later that

$$N_n \to \sqrt{\frac{2}{l_1}}.$$

First of all, let us show that all the functions $v_n(s)$ are bounded by some constant that is independent of n. Let us denote $\max|v_n(s)|$ for $0 \leqslant s \leqslant l_1$ by M_n. Then, from equation (22.26), we have

$$|v_n(s)| \leqslant 1 + \frac{1}{\sqrt{\lambda_n}} M_n \int_0^{l_1} |R(\tau)|\, d\tau$$

and, consequently,

$$M_n \leqslant 1 + \frac{M_n}{\sqrt{\lambda_n}} \int_0^{l_1} |R(\tau)|\, d\tau,$$

$$M_n \leqslant \frac{1}{1 - \frac{1}{\sqrt{\lambda_n}} \int_0^{l_1} |R(\tau)|\, d\tau} = 1 + O\left(\frac{1}{\sqrt{\lambda_n}}\right). \tag{22.27}$$

Since $\lambda_n \to \infty$ as $n \to \infty$ (see subsection 7), this inequality proves the boundedness of the functions $v_n(s)$.

In what follows, we shall need an analogous estimate for

the first and second derivatives of the eigenfunctions. To get it, we differentiate the integral equation (22.26). We obtain

$$v_n'(s) = \sqrt{\lambda_n}\cos\sqrt{\lambda_n}\,s + \int_0^s R(\tau)\,v_n(\tau)\cos\sqrt{\lambda_n}(s-\tau)\,d\tau,$$

$$v_n''(s) = -\lambda_n\sin\sqrt{\lambda_n}\,s - \sqrt{\lambda_n}\int_0^s R(\tau)\,v_n(\tau)\sin\sqrt{\lambda_n}(s-\tau)\,d\tau + R(s)\,v_n(s),$$

so that

$$|v_n'(s)| \leqslant \sqrt{\lambda_n} + O(1), \quad |v_n''(s)| \leqslant \lambda_n + O(\sqrt{\lambda_n}). \tag{22.28}$$

Let us now compute

$$\int_0^{l_1} v_n^2(s)\,ds,$$

that is, let us find the factor by which the functions $v_n(s)$ (and consequently their derivatives) differ from the normalised eigenfunctions $u_n(s)$ and their respective derivatives). From (22.26), we have

$$v_n^2(s) = \sin^2\sqrt{\lambda_n}\,s + O\left(\frac{1}{\sqrt{\lambda_n}}\right).$$

From this we get

$$\int_0^{l_1} v_n^2(s)\,ds = \frac{l_1}{2} - \frac{\sin 2\sqrt{\lambda_n}\,l_1}{4\sqrt{\lambda_n}} + O\left(\frac{1}{\sqrt{\lambda_n}}\right) = \frac{l_1}{2} + O\left(\frac{1}{\sqrt{\lambda_n}}\right).$$

Therefore, for $u_n(s)$, we immediately obtain estimates analogous to (22.27) and (22.28):

$$\left.\begin{aligned} |u_n(s)| &\leqslant \sqrt{\frac{2}{l_1}} + O\left(\frac{1}{\sqrt{\lambda_n}}\right), \\ |u_n'(s)| &\leqslant \sqrt{\lambda_n}\sqrt{\frac{2}{l_1}} + O(1), \\ |u_n''(s)| &\leqslant \lambda_n\sqrt{\frac{2}{l_1}} + O(\sqrt{\lambda_n}). \end{aligned}\right\} \tag{22.29}$$

By a change of variables indicated in formulae (22.22), we immediately obtain from (22.29) the corresponding results

$$f_N(x) = f(x) - \sum_{n=1}^{N} c_n X_n(x), \quad \delta_N^2 = \int_0^l \rho(x) f_N^2(x)\,dx,$$

$$\varphi_N(x) = \frac{f_N(x)}{\delta_N}.$$

We need to show that $\delta_N^2 \to 0$ as $N \to \infty$. Since

$$\int_0^l \rho\,\varphi_N^2(x)\,dx = 1$$

and since, in addition,

$$\int_0^l \rho\,\varphi_N(x) X_n(x)\,dx = 0 \quad (n = 1, \ldots, N),$$

the function $\varphi_N(x)$ is one of the admissible functions of the variational problem examined in subsection 3 of this section*. The value of the minimum of $D(X)$ for this problem is equal to λ_{N+1}. Consequently,

$$D(\varphi_N) \geqslant \lambda_{N+1}.$$

Let us now calculate $D(\varphi_N)$. Using the notations of subsection 3, we find

$$\begin{aligned} D(\varphi_N) &= \int_0^l (p\varphi_N'^2 + q\varphi_N^2)\,dx = \frac{1}{\delta_N^2}\int_0^l (pf_N'^2 + qf_N^2)\,dx \\ &= \frac{1}{\delta_N^2}\int_0^l \Big[p\Big(f' - \sum_{n=1}^{N} c_n X_n'\Big)^2 + q\Big(f - \sum_{n=1}^{N} c_n X_n\Big)^2\Big]\,dx \\ &= \frac{1}{\delta_N^2}\Big[D(f) - 2\sum_{n=1}^{N} c_n D(f, X_n) + \sum_{n=1}^{N}\sum_{m=1}^{N} c_n c_m D(X_n, X_m)\Big] \\ &\geqslant \lambda_{N+1}. \end{aligned} \tag{22.32}$$

On the basis of Theorem 3 of subsection 3, we have

$$D(f, X_n) = \lambda_n c_n, \quad D(X_n, X_n) = D(X_n) = \lambda_n,$$
$$D(X_m, X_n) = 0 \text{ when } n \neq m.$$

When we substitute these values of the functionals into (22.32) we obtain

* See footnote on page 170.

$$\frac{1}{\delta_N^2}[D(f) - \sum_{n=1}^{N} \lambda_n c_n^2] \geqslant \lambda_{N+1},$$

so that

$$\delta_N^2 \leqslant \frac{D(f) - \sum_{n=1}^{N} \lambda_n c_n^2}{\lambda_{N+1}}. \tag{22.33}$$

In accordance with (22.19) there exist only a finite number of negative λ_n. Therefore, the numerator on the right-hand side of (22.33) is bounded for all N. Since $\lambda_{N+1} \to \infty$ as $N \to \infty$, it follows that $\delta_N^2 \to 0$ as $N \to \infty$. Thus, the series (22.30) converges in mean for every differentiable function that vanishes at the point $x=0$ and $x=l$.

To free ourselves from the restrictions imposed on $f(x)$ let us note that, for every piecewise-continuous square-integrable function $f(x)$, there exists a continuous differentiable function $f^*(x)$ that vanishes at the end points of the interval $[0, l]$ satisfying the inequality

$$\int_0^l \rho [f(x) - f^*(x)]^2 dx < \varepsilon_1,$$

where ε_1 is an arbitrary positive number.

Suppose, further, that N is chosen sufficiently large that

$$\int_0^l \rho(x) [f^*(x) - \sum_{n=1}^{N} c_n^* X_n(x)]^2 dx < \varepsilon_2;$$

where the c_n^* are the Fourier coefficients for $f^*(x)$. Then,

$$\int_0^l \rho(x) [f(x) - \sum_{n=1}^{N} c_n X_n(x)]^2 dx$$

$$\leqslant \int_0^l \rho(x) [\,|f(x) - f^*(x)| + |f^*(x) - \sum_{n=1}^{N} c_n^* X_n(x)|\,]^2 dx$$

$$\leqslant \varepsilon_1 + \varepsilon_2 + 2\int_0^l \rho(x) |f(x) - f^*(x)|\,|f^*(x)$$

$$- \sum_{n=1}^{N} c_n^* X_n(x)|\, dx \leqslant \varepsilon_1 + \varepsilon_2 + 2\sqrt{\varepsilon_1 \cdot \varepsilon_2}.$$

To get an estimate of this last integral, we used Bunyakovskii's inequality.

We have thus shown that, for every square-integrable function $f(x)$, there exists an N and constants c_n such that the quantity

$$\int_0^l \rho(x)\,[f(x) - \sum_{n=1}^{N} c_n X_n(x)]^2 dx \tag{22.34}$$

will be arbitrarily small. However, we know (see, for example, my Letsii po teorii integral'nykh uravnenii (Lectures on the theory of integral equations), Gostekhizdat, 1951, pp. 66-67), an integral of the form (22.34) has its smallest value when the c_n are the Fourier coefficients for the function $f(x)$. Therefore, if we take these coefficients for the c_n in (22.34), the value of the integral (22.34) will not increase.

By using the orthogonality of the functions $X_n(x)$, it is easy to show that

$$\delta_N^2 = \int_0^l \rho\,[f(x)]^2\,dx - \sum_{n=1}^{N} c_n^2 \geqslant 0$$

(Bessel's inequality), where the c_n are the Fourier coefficients for the function $f(x)$. Consequently, the condition that the system of functions is complete may be written in the form of the following equation:

$$\sum_{n=1}^{\infty} c_n^2 = \int_0^l \rho\,[f(x)]^2\,dx \tag{22.35}$$

(Parseval's equation).

10. Let us now prove the following fundamental theorem.

Theorem. Suppose that a continuously differentiable function on the interval $[0, l]$ vanishes at the end points of that interval. Then, its 'Fourier series' (22.30) converges absolutely and uniformly to that function.

It will be sufficient to show that this series converges absolutely and uniformly in general. Indeed, since this series converges 'in mean' to $f(x)$, if it converges uniformly its

limit cannot be any other function.

Suppose that n_0 is sufficiently large that $\lambda_n > 0$ for $n \geqslant n_0$. By using Cauchy's inequality, we can write, for $n \geqslant n_0$

$$\sum_{k=n}^{n+s} |c_k X_k| = \sum_{k=n}^{n+s} |c_k \sqrt{\lambda_k}| \cdot \left|\frac{X_k}{\sqrt{\lambda_k}}\right|$$
$$\leqslant \sqrt{\sum_{k=n}^{n+s} c_k^2 \lambda_k} \sqrt{\sum_{k=n}^{n+s} \frac{X_k^2}{\lambda_k}} \leqslant \sqrt{\sum_{k=n_0}^{n+s} c_k^2 \lambda_k} \sqrt{\sum_{k=n}^{n+s} \frac{X_k^2}{\lambda_k}}.$$

Let us now apply inequality (22.33) for an estimate of the first factor and let us take the quantity

$$\max_{k,\, x} |X_k(x)| = M$$

from under the radical. Since it follows from inequality (22.33) that

$$\sum_{k=n_0}^{\infty} c_k^2 \lambda_k \leqslant D(f) + \sum_{k=1}^{n_0} c_k^2 |\lambda_k| \leqslant M_1^2,$$

we have

$$\sum_{k=n}^{n+s} |c_k X_k| \leqslant M_1 M \sqrt{\sum_{k=n}^{n+s} \frac{1}{\lambda_k}}.$$

Since, by virtue of (22.19),

$$\lambda_k \geqslant c_1 k^2 + c_2,$$

it follows that

$$\frac{1}{\lambda_k} \leqslant \frac{1}{c_1 k^2 + c_2},$$

and the series

$$\sum_{k=n_0}^{\infty} \frac{1}{\lambda_k}$$

converge. Consequently, for arbitrary positive ε and s and sufficiently large n,

$$\sum_{k=n}^{n+s} \frac{1}{\lambda_k} < \frac{\varepsilon^2}{M^2 M_1^2}$$

and hence,

$$\sum_{k=n}^{n+s} |c_k X_k| < \varepsilon,$$

that is, the series

$$\sum c_k X_k(x)$$

converges absolutely and uniformly.

23. JUSTIFICATION OF THE FOURIER METHOD

1. Consider equation (21.1). Let us assume that the coefficients in this equation are three times continuously differentiable functions in the cylinder $\overline{C}_T$, that $A(t) > a_0 > 0$, and that $C(x) < c_0 < 0$, in other words, equation (21.1) is a hyperbolic equation*.

Let us seek a solution of equation (21.1) that is twice continuously differentiable in $\overline{C}_T$ and that satisfies the initial conditions

$$u(0, x) = \varphi_0(x), \quad u_t'(0, x) = \varphi_1(x) \tag{23.1}$$

and the boundary conditions

$$u(t, 0) = u(t, l) = 0. \tag{23.2}$$

The Fourier method leads to an examination of the series (21.12) (see Section 21). The functions $X_k(x)$ are eigenfunctions of equation (22.1). Suppose that

$$L(f) \equiv (pf')' - qf.$$

Then, equation (22.1) can be written as follows:

$$L(X_k) = -\lambda_k \rho X_k.$$

Theorem. If $\varphi_0(x)$ has a continuous third derivative on the interval $[0, l]$ and satisfies the conditions

$$\varphi_0 = L(\varphi_0) = 0 \quad \text{at} \quad x = 0 \text{ and } x = l, \tag{23.3}$$

* It is easy to show that all of the theorems in Section 22 and the fundamental theorem in Section 23 are valid if $C(x)$ and $A(t)$ are twice continuously differentiable and if $D(t)$, $E(x)$, and $F_2(x)$ have continuous first-order derivatives and $F_1(t)$ is continuous.

and if $\varphi_1(x)$ has a continuous second derivative on that interval and satisfies the conditions

$$\varphi_1 = 0 \quad \text{at} \quad x = 0 \text{ and } x = l, \tag{23.4}$$

then the function $u(t, x)$ defined by the series (21.12) has continuous second-order derivatives and satisfies equation (21.1), the initial conditions (23.1), and the boundary conditions (23.2) throughout $\bar{C}_T$. Here, term-wise twice differentiability of the series (21.12) with respect to t and x is possible. The series that result from these differentiations converge absolutely and uniformly in $\bar{C}_T$ †.

Proof: Consider the series (21.12) constructed in Section 21

$$u(x, t) = \sum_{k=1}^{\infty} X_k(x) \left[A_k T_k^*(t) + B_k T_k^{**}(t)\right]. \tag{23.5}$$

Here,

$$L(X_k) = -\lambda_k \rho X_k, \quad \int_0^l \rho X_k^2(x)\, dx = 1,$$

$$A_k = \int_0^l \rho \varphi_0 X_k\, dx \text{ and } B_k = \int_0^l \rho \varphi_1 X_k\, dx.$$

The functions T_k^* and T_k^{**} are solutions of equation (21.6) at $\lambda = \lambda_k$ and they satisfy the initial conditions

$$T_k^*(0) = 1, \quad \frac{dT_k^*(0)}{dt} = 0,$$

$$T_k^{**}(0) = 0, \quad \frac{dT_k^{**}(0)}{dt} = 1.$$

† Conditions (23.3) and (23.4) are necessary conditions for the existence in $\bar{C}_T$ of a twice continuously differentiable solution of the problem stated. To see this, note that condition (23.2) implies that $\partial u/\partial t$ and $\partial^2 u/\partial t^2$ are equal to zero at $x = 0$ and at $x = l$. In view of this fact, we obtain from equation (21.1) that, at $x = 0$ and $x = l$

$$C(x) \frac{\partial^2 u}{\partial x^2} + E(x) \frac{\partial u}{\partial x} + F_2(x) u = 0,$$

that is $L(\phi_0) = 0$ at $x = 0$ and $x = l$.

By a change of variables analogous to that of (22.20), we can reduce equation (21.6) to the form

$$w'' + \lambda_k w = R(s)\,w. \tag{23.6}$$

Since here $T(t) = \psi(t)\,w$, where $\psi(t)$ is some function that is independent of k, the functions $\overset{*}{w}_k$ and $\overset{**}{w}_k$ corresponding to the functions $\overset{*}{T}_k$ and $\overset{**}{T}_k$ satisfy the initial conditions

$$\overset{*}{w}_k(0) = a^*, \quad \overset{*}{w}_k{}'(0) = b^* \text{ and } \overset{**}{w}_k(0) = 0, \quad \overset{**}{w}_k{}'(0) = b^{**},$$

where a^*, b^*, and b^{**} are numbers that are independent of k. For solutions of equation (23.6), we may write an integral equation of the form (22.25). By a procedure analogous to that followed in Section 22, we may use this integral equation to obtain estimates for $\overset{*}{w}_k$ and $\overset{**}{w}_k$ and their derivatives. Here, we obtain the following estimates for the functions $T^*(t)$ and $T^{**}(t)$ for sufficiently large k:

$$\left.\begin{aligned} &|\overset{*}{T}_k| < M, \quad \left|\frac{d\overset{*}{T}_k}{dt}\right| < M\sqrt{\lambda_k}, \quad \left|\frac{d^2\overset{*}{T}_k}{dt^2}\right| < M\lambda_k, \\ &|T^{**}| < \frac{M}{\sqrt{\lambda_k}}, \quad \left|\frac{d\overset{**}{T}_k}{dt}\right| < M, \quad \left|\frac{d^2\overset{**}{T}_k}{dt^2}\right| < M\sqrt{\lambda_k}, \end{aligned}\right\} \tag{23:7}$$

where $M > 0$ is some constant.

Now, let us estimate the Fourier coefficients of the function $\varphi_0(x)$:

$$A_k = \int_0^l \rho\varphi_0 X_k\,dx = -\int_0^l \varphi_0 \frac{L(X_k)}{\lambda_k}\,dx$$
$$= -\int_0^l \frac{1}{\lambda_k} L(\varphi_0)\,X_k\,dx. \tag{23.8}$$

We obtained this last equation by twice integrating by parts and using the boundary conditions

$$\varphi_0(0) = \varphi_0(l) = X_k(0) = X_k(l) = 0.$$

We obtain from equation (23.8)

$$\lambda_k A_k = -\int_0^l \rho \frac{L(\varphi_0)}{\rho} X_k\,dx,$$

that is, the $\lambda_k A_k$ are the Fourier coefficients of the continuously differentiable function $H(x) = -\frac{L(\varphi_0)}{\rho}$ which satisfies the conditions $H(0) = H(l) = 0$.

Uniform convergence of the series

$$\sum_{k=1}^{\infty} |\lambda_k A_k| \, |X_k|$$

follows from the theorem of subsection 10 of Section 22. From inequality (22.33), we easily see that the series

$$\sum_{k=1}^{\infty} \lambda_k^3 A_k^2$$

converges.

Let us now estimate $B_k = \int_0^l \rho\varphi_1 X_k \, dx$. When we again use equation (22.1), integrate by parts twice, and use the boundary conditions

$$\varphi_1(0) = \varphi_1(l) = X_k(0) = X_k(l) = 0,$$

we obtain

$$B_k = \int_0^l \rho\varphi_1 X_k \, dx = -\int_0^l \frac{1}{\lambda_k} \rho \frac{L(\varphi_1)}{\rho} X_k \, dx = \frac{\beta_k}{\lambda_k},$$

where the β_k are the Fourier coefficients of the continuous function $H_1(x) = -\frac{L(\varphi_1)}{\rho}$. On the basis of equation (22.35), we have

$$\sum_{k=1}^{\infty} \beta_k^2 = \int_0^l \rho \left(\frac{L(\varphi_1)}{\rho} \right)^2 dx.$$

By using the estimates (23.7) and (22.29) and by taking account of (22.20), we can easily show that the numerical series

$$\sum_{k=1}^{\infty} \left(|\lambda_k| \, |A_k| + |B_k| \sqrt{|\lambda_k|} \right), \tag{23.9}$$

implies absolute and uniform convergence of the series

(21.12) and of the series obtained from it by twice differentiating it termwise with respect to x and t, since, for sufficiently large values of k, the terms of these series do not exceed in absolute value the terms of the series

$$M_1 \sum_{k=1}^{\infty} (|\lambda_k| |A_k| + |B_k| \sqrt{|\lambda_k|}),$$

where M_1 is some positive constant. To show that the series (23.9), converges, we note that, for sufficiently large n,

$$\sum_{k=n}^{n+m} (|\lambda_k| |A_k| + |B_k| \sqrt{\lambda_k}) = \sum_{k=n}^{n+m} \left(|A_k| |\lambda_k|^{3/2} \frac{1}{\sqrt{\lambda_k}} + \frac{|\beta_k|}{\sqrt{\lambda_k}} \right)$$
$$\leqslant \sqrt{\sum_{k=n}^{n+m} A_k^2 \lambda_k^3} \sqrt{\sum_{k=n}^{n+m} \frac{1}{\lambda_k}} + \sqrt{\sum_{k=n}^{n+m} \beta_k^2} \sqrt{\sum_{k=n}^{n+m} \frac{1}{\lambda_k}}. \quad (23.10)$$

Here, we used Cauchy's inequality. Convergence of the series

$$\sum_{k=1}^{\infty} A_k^2 \lambda_k^3, \quad \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \quad \text{and} \quad \sum_{k=1}^{\infty} \beta_k^2$$

and inequality (23.10) imply convergence of the series (23.9). This completes the proof of the theorem.

2. Let us now show that the next problem for a hyperbolic equation of type (21.1) has a unique solution. We have already shown, in Section 18, that the mixed problem for the wave equation has a unique solution.

By integrating by parts, one can easily show that for any two functions $u(t, x)$ and $v(t, x)$ that are twice continuously differentiable in $\bar{C}_T$ for $0 < T_1 \leqslant T$, the following formula holds

$$\iint_{\bar{U}_{T_1}} \Big\{ v \Big[A(t) \frac{\partial^2 u}{\partial t^2} + C(x) \frac{\partial^2 u}{\partial x^2} + D(t) \frac{\partial u}{\partial t} + E(x) \frac{\partial u}{\partial x}$$
$$+ (F_1(t) + F_2(x)) u \Big] - u \Big[\frac{\partial^2 (A(t) v)}{\partial t^2} + \frac{\partial^2 (C(x) v)}{\partial x^2}$$
$$- \frac{\partial (D(t) v)}{\partial t} - \frac{\partial (E(x) v)}{\partial x} + (F_1(t) + F_2(x)) v \Big] \Big\} \, dx \, dt$$

(continued over page)

$$=\int_0^l \left[vA(t)\frac{\partial u}{\partial t}-u\frac{\partial (Av)}{\partial t}+Duv\right]_{t=T_1} dx$$
$$-\int_0^l \left[vA(t)\frac{\partial u}{\partial t}-u\frac{\partial (Av)}{\partial t}+Duv\right]_{t=0} dx$$
$$+\int_0^{T_1} \left[vC(x)\frac{\partial u}{\partial x}-u\frac{\partial (Cv)}{\partial x}+Euv\right]_{x=l} dt$$
$$-\int_0^{T_1} \left[vC(x)\frac{\partial u}{\partial x}-u\frac{\partial (Cv)}{\partial x}+Euv\right]_{x=0} dt. \qquad (23.11)$$

Suppose that $u(t, x)$ satisfies equation (23.1) and the conditions

$$u(0, x)=0, \quad u_t'(0, x)=0, \quad u(t, 0)=0, \quad u(t, l)=0 \quad (23.12)$$

in $\bar{C}_T$. Let us show that $u(t, x)\equiv 0$.

Suppose otherwise. Assume that $u(t, x)$ is nonzero at the point (T_1, x_1). Let us apply formula (23.11) to the function $u(t, x)$ and to the function $v(t, x)$, which we choose in such a way that, throughout $\bar{C}_{T_1}$, it will satisfy the equation

$$\frac{\partial^2 (A(t)v)}{\partial t^2}+\frac{\partial^2 (C(x)v)}{\partial x^2}-\frac{\partial (D(t)v)}{\partial t}-\frac{\partial (E(x)v)}{\partial x} + (F_1(t)+F_2(x))v=0 \qquad (23.13)$$

and the conditions

$$v(t, 0)=0, \quad v(t, l)=0, \quad v(T_1, x)=0, \quad v_t'(T_1, x)=\alpha(x), \quad (23.14)$$

where $\alpha(x)$ is a smooth nonnegative function that is nonzero only in a small neighbourhood of the point (T_1, x_1) in which $u(t, x)$ keeps its sign. The existence of the function $v(t, x)$ follows from the preceding theorem since the solution of (23.13) is of the form (21.1).

It is easy to see on the basis of relations (21.1), (23.12), (23.13), and (23.14) that the left member of equation (23.11) is equal to zero and that the right member is equal to

$$\int_0^l -u(T_1, x)A(T_1)\alpha(x)\,dx\neq 0.$$

This contradiction shows that $u\equiv 0$.

Problem. Show that the solution of the mixed problem for

equation (21.1) depends continuously on the initial conditions, that is, that the solution $u(t, x)$ of equation (21.1) satisfying conditions (23.1) and (23.2) will be arbitrarily small in absolute value throughout $\overline{C}_T$ if $|\varphi_0(x)|$, $\left|\frac{\partial\varphi_0}{\partial x}\right|$ and $|\varphi_1(x)|$ is sufficiently small for all x in the interval $[0, l]$.

To prove this assertion, one must use the estimates (23.7) and (22.29), equation (22.35) for the function $\varphi_1(x)$, inequality (22.33) for the function $\varphi_0(x)$, and the convergence of the series $\sum_{k=1}^{\infty}\frac{1}{\lambda_k}$.

Remark: It is easy to show that, if $u(t, x)$ satisfies equations (21.1), the initial conditions (23.1), and the boundary conditions (23.2) in $\overline{C}_T$, the integral

$$\iint_{U_T} \rho u^2(t, x)\,dx\,dt$$

will be arbitrarily small if $\int_0^l \rho\varphi_0^2(x)\,dx$ and $\int^l \rho\varphi_1^2(x)\,dx$ are sufficiently small.

To show this, we represent $u(t, x)$ in the form of a series (21.12) and obtain

$$\begin{aligned}\iint_{U_T}&\rho u^2(t, x)\,dx\,dt\\ &=\iint_{U_T}\rho\left|\sum_{k=1}^{\infty}X_k(x)\,A_kT_k^*(t)+\sum_{k=1}^{\infty}X_k(x)\,B_kT_k^{**}(t)\right|^2dx\,dt\\ &\qquad\leqslant 2\iint_{U_T}\rho\Big(\sum_{k=1}^{\infty}X_k(x)\,A_kT_k^*(t)\Big)^2dx\,dt\\ &\qquad+2\iint_{U_T}\rho\Big(\sum_{k=1}^{\infty}X_k(x)\,B_kT_k^{**}(t)\Big)^2dx\,dt\\ &\leqslant K_1\sum_{k=1}^{\infty}A_k^2+K_2\sum_{k=1}^{\infty}B_k^2=K_1\int_0^1\rho\varphi_0^2(x)\,dx+K_2\int_0^l\rho\varphi_1^2(x)\,dx,\end{aligned}$$

where K_1 and K_2 are positive constants that are independent of φ_0 and φ_1. In the derivation of these estimates, we used the elementary inequality $(a+b)^2\leqslant 2a^2+2b^2$, the orthogonality (with weight $\rho(x)$) of the eigenfunctions (which are as-

sumed to be normalised), the boundedness of the functions T_k^* and T_k^{**}, and Parseval's equation (22.35).

3. If the initial functions $\varphi_0(x)$ and $\varphi_1(x)$ do not satisfy the conditions formulated in the theorem of the present section, there may not be any solution to the mixed problem for equation (21.1) that is twice continuously differentiable in $\overline{C}_T$. However, if $\varphi_0(x)$ is a continuously differentiable function that vanishes at $x=0$ and $x=l$ and if $\varphi_1(x)$ is a continuous function on the interval $[0, l]$, then the series (21.12) converges and defines in $\overline{C}_T$ a continuous function $u(t, x)$. The function $u(t, x)$ will then be a generalised solution of the mixed problem for equation (21.1) corresponding to the initial conditions (23.1) and the boundary conditions (23.2).

We call the function $u(t, x)$ a generalised solution of equation (21.1) with initial conditions (23.1) and boundary conditions (23.2) if $u(t, x)$ is the limit in $\overline{C}_T$ as $n \to \infty$ of a uniformly convergent sequence $u_n(t, x)$ of solutions of equation (21.1) with boundary conditions (23.2) and initial conditions

$$\left.\begin{aligned} u_n(0, x) &= \varphi_0^n(x), \\ \frac{\partial u_n(0, x)}{\partial t} &= \varphi_1^n(x), \end{aligned}\right\} \tag{23.15}$$

and if

$$\int_0^l \rho\,[\varphi_0(x) - \varphi_0^n(x)]^2\,dx \to 0$$

and

$$\int_0^l \rho\,[\varphi_1(x) - \varphi_1^n(x)]^2\,dx \to 0 \tag{23.16}$$

as $n \to \infty$.

Let us show that if $\varphi_0(x)$ is a continuously differentiable function that vanishes at $x=0$ and $x=l$ and if $\varphi_1(x)$ is a continuous function of $[0, l]$, then to equation (21.1) with conditions (23.1) and (23.2) there corresponds a unique generalised solution. The existence of a generalised solution follows from the fact that the partial sums of the series (21.12) form a sequence $u_n(t, x)$, that satisfies the required

conditions, so that, consequently, the series (21.12) is a generalised solution. Let us now show that the generalised solution is unique.

Suppose that to two distinct sequences of functions $\varphi_0^n(x)$, $\varphi_1^n(x)$ and $\tilde{\varphi}_0^n(x)$, $\tilde{\varphi}_1^n(x)$ there corresponded two distinct limiting functions $u(t,x)$ and $\tilde{u}(t,x)$ for the sequences $u_n(t,x)$ and $\tilde{u}_n(t,x)$ We would then have

$$\iint_{C_T} \rho(u-\tilde{u})^2\,dx\,dt$$
$$=\iint_{C_T} \rho[(u-u_n)+(u_n-\tilde{u}_n)+(\tilde{u}_n-\tilde{u})]^2\,dx\,dt$$
$$\leqslant 3\iint_{C_T} \rho(u-u_n)^2\,dx\,dt+3\iint_{C_T} \rho(u_n-\tilde{u}_n)^2\,dx\,dt$$
$$+3\iint_{C_T} \rho(\tilde{u}_n-\tilde{u})^2\,dx\,dt. \tag{23.17}$$

On the basis of the remark in subsection 2 of the preceding section, the integral

$$\iint_{C_T} \rho(u_n-\tilde{u}_n)^2\,dx\,dt$$

approaches zero as $n\to\infty$ since

$$\int_0^l \rho(\varphi_0^n-\tilde{\varphi}_0^n)^2\,dx \quad \text{and} \quad \int_0^l \rho(\varphi_1^n-\tilde{\varphi}_1^n)^2\,dx$$

approach 0 as $n\to\infty$. Since the two integrals on the right-hand side of inequality (23.17) also approach zero, we have

$$\iint_{C_T} \rho(u-\tilde{u})^2\,dx\,dt=0.$$

Since $u-\tilde{u}$ and $\rho>0$ are continuous functions, we have $u(t,x)\equiv\tilde{u}(t,x)$.

From the definition of a generalised solution to the mixed problem for equation (21.1) it follows that if for $\varphi_0(x)$ and $\varphi_1(x)$, there exists a solution to the mixed problem that is twice continuously differentiable in $\overline{C}_T$, the generalised solution of the mixed problem coincides with this solution.

A function $u(t,x)$ such that

$$\lim_{n \to \infty} \iint_{U_T} \rho (u_n - u)^2 \, dx \, dt = 0, \tag{23.18}$$

where the functions $u_n(t, x)$ are solutions to equation (21.1) under the boundary conditions (23.2) and the initial conditions (23.15), and such that the relations (23.16) are satisfied is sometimes called a generalised solution of the mixed problem for equation (21.1).

Let us show other possible definitions of a generalised solution of the mixed problem in which the integral identities (cf. Section 9) are used. For simplicity, let us consider an equation of the form

$$P(u) \equiv \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x}\left(p(x)\frac{\partial u}{\partial x}\right) + q(x)\, u = 0. \tag{23.19}$$

A function $u(t, x)$ that is continuously differentiable in $\overline{C}_T$ and that satisfies the conditions

$$u(0, x) = \varphi_0(x), \quad u(t, 0) = u(t, l) = 0 \tag{23.20}$$

and the integral identity

$$\iint_{U_T} \left(\frac{\partial u}{\partial t}\frac{\partial \sigma}{\partial t} - p\frac{\partial u}{\partial x}\frac{\partial \sigma}{\partial x} - qu\sigma\right) dx\, dt + \int_0^l \varphi_1(x)\, \sigma(0, x)\, dx = 0 \tag{23.21}$$

for an arbitrary continuously differentiable function $\sigma(t, x)$ that vanishes at $t = T$, at $x = 0$, and $x = l$ is called a generalised solution of the mixed problem for equation (23.19) under the initial and boundary conditions (23.1)-(23.2).

Sometimes, it is convenient to use the following definition. A function $u(t, x)$ satisfying the integral identity

$$\iint_{U_T} uP(\sigma)\, dx\, dt + \int_0^l \varphi_0(x)\frac{\partial \sigma}{\partial t}(0, x)\, dx - \int_0^l \varphi_1(x)\, \sigma(0, x)\, dx = 0, \tag{23.22}$$

where $\sigma(t, x)$ is an arbitrary twice continuously differentiable function such that

$$\sigma(t, 0) = \sigma(t, l) = \sigma(T, x) = \frac{\partial \sigma}{\partial t}(T, x) = 0 \qquad (23.23)$$

is called a generalised solution of the mixed problem for equation (23.19) under the conditions (23.1)-(23.2).

Obviously, a generalised solution defined by the identity (23.21) under the conditions (23.20) will at the same time be a generalised solution in the sense of (23.22). The converse does not in general hold.

By introducing generalised solutions, we can in some degree, broaden the class of initial conditions under which a solution of the mixed problem exists. Here, it is very important that the uniqueness theorem remain valid in the new class of solutions.

Problem 1. Show that the generalised solution of equation (23.19) under the conditions (23.1)-(23.2) that is defined by the relation (23.18) (where $\rho = 1$) exists and is unique if the functions $\varphi_0(x)$ and $\varphi_1(x)$ are piecewise continuous and square-integrable on the interval $[0, l]$.

2. Show that the generalised solution of equation (23.19) under the conditions (23.1)-(23.2) that is defined by the relations (23.20)-(23.21) exists and is unique if the function $\varphi_0(x)$ is twice continuously differentiable on the interval $[0, l]$ and the function $\varphi_1(x)$ is at least once continuously differentiable on that interval and if

$$\varphi_0(0) = \varphi_0(l) = \varphi_1(0) = \varphi_1(l) = 0 \text{ and } q(x) \geqslant 0.$$

Hint: To prove the uniqueness, use the function

$$\sigma(t, x) = \int_T^t [u_1(\tau, x) - u_2(\tau, x)]\, d\tau,$$

where u_1 and u_2 are two generalised solutions of the same mixed problem.

3. Show that the generalised solution of equation (23.19) under the conditions (23.1)-(23.2) that is defined by the relation (23.22) exists and is unique if the function $\varphi_0(x)$ is continuous on the interval $[0, l]$, vanishes at $x = 0$ and $x = l$, and has a piecewise continuous square-integrable derivative on that interval and if the function $\varphi_1(x)$ is piecewise continuous and square-integrable on the same interval $[0, l]$.

Hint: To prove the uniqueness, use the results of subsection 4 of the present section.

4. The Fourier method for a nonhomogeneous hyperbolic equation Let us consider in C_T the mixed problem for the equation

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial}{\partial x}\left(p(x)\frac{\partial u}{\partial x}\right)-q(x)u+f(t, x)\equiv L(u)+f(t, x), \quad (23.24)$$

That is, let us seek a solution to this equation that is twice continuously differentiable in $\overline{C}_T$ that satisfies the initial conditions

$$u(0, x)=\varphi_0(x), \quad u'_t(0, x)=\varphi_1(x) \qquad (23.25)$$

and the boundary conditions

$$u(t, 0)=u(t, l)=0. \qquad (23.26)$$

Here, it will be sufficient for us to find a solution satisfying the conditions (23.25) when $\varphi_0(x)=\varphi_1(x)\equiv 0$ since we will obtain the desired solution by adding to it the series (21.12).

Let us seek a solution $u(t, x)$ of this problem in the form of a Fourier series

$$\sum_{k=1}^{\infty} a_k(t)X_k(x)$$

of the eigenfunctions of equation $L(X)=-\lambda X$ with the boundary conditions $X(0)=X(l)=0$. When we expand the function $f(t, x)$ in a Fourier series of these eigenfunctions and equate the Fourier coefficients on the right and left sides of equation (23.24), we shall obtain differential equations for determining the Fourier coefficients $a_k(t)$ of the form

$$a''_k(t)=-\lambda_k a_k(t)+f_k(t), \qquad (23.27)$$

where

$$f_k(t)=\int_0^l f(t, x)X_k(x)\,dx$$

and

$$L(X_k)=-\lambda_k X_k.$$

It is easy to show that the function

$$\frac{1}{\sqrt{\lambda_k}}\int_0^t f_k(\tau)\sin\sqrt{\lambda_k}(t-\tau)\,d\tau.$$

is a solution of equation (23.27) satisfying the conditions

$$a_k(0) = a'_k(0) = 0.$$

Thus, the solution $u(t, x)$ of equation (23.24) that satisfies conditions (23.26) and the conditions

$$u(0, x) = u_t(0, x) = 0, \tag{23.28}$$

must be expressed in the form of the series

$$u(t, x) = \sum_{k=1}^{\infty} \left(\frac{1}{\sqrt{\lambda_k}} \int_0^t f_k(\tau) \sin \sqrt{\lambda_k}(t-\tau)\, d\tau \right) X_k(x). \tag{23.29}$$

If this series and the series obtained from it by differentiating it termwise twice with respect to x and t converge uniformly in $\overline{C}_T$, then the sum of this series is a function that is twice continuously differentiable in $\overline{C}_T$ and that satisfies equation (23.24) and the conditions (23.26) and (23.28). Convergence of these series will be ensured if we require that the continuous function $f(t, x)$ have a continuous second derivative with respect to x and that the conditions $f(t, 0) = f(t, l) = 0$ be satisfied for all t. Here, we assume that the coefficients $p(x)$ and $q(x)$ have continuous second derivatives. Proof of this is exactly analogous to the proof of the fundamental theorem of the present section. The Fourier coefficients $f_k(t)$ of the function $f(t, x)$ are evaluated in the same way as the coefficients B_k of the series (21.12).

24. APPLICATION OF GREEN'S FUNCTION TO THE EIGENVALUE PROBLEM AND TO THE JUSTIFICATION OF THE FOURIER METHOD

The existence of a complete system of eigenfunctions in an eigenvalue problem and the basic properties of this system can be shown without solving the variational problems but by a completely different method. This can be done by reducing the boundary-value problem to a Fredholm integral equation of the second kind. This reduction can be carried out with the aid of what is called a Green's function. We now turn to the construction of this function.

1. Consider the problem of finding the solution on the interval $(0, l)$ of the equation

$$(pX')' - qX = f(x), \tag{24.1}$$

satisfying the conditions

$$X(0) = X(l) = 0. \tag{24.2}$$

Together with equation (24.1), let us consider the equation

$$(pY')' - qY = g_\varepsilon(x, x_0) \tag{24.3}$$

with the same left member as equation (24.1) but with the free term

$$g_\varepsilon(x, x_0) = \begin{cases} \frac{1}{\varepsilon} & \text{for } x_0 - \frac{\varepsilon}{2} < x < x_0 + \frac{\varepsilon}{2}, \\ 0 & \text{for all other } x. \end{cases}$$

Here ε and x_0 are constants such that $\varepsilon > 0; 0 < x_0 < l$, and $0 < \frac{\varepsilon}{2} \leqslant \min\{x_0, l - x_0\}$. Let us assume that we know the solution $Y_\varepsilon(x, x_0)$ of this equation that satisfies the same boundary conditions (24.2) but which depends on the constants ε and x_0 *.

Let us multiply equation (24.1) by Y_ε. Then, let us replace Y in equation (24.3) with Y_ε and multiply this equation by X. Finally, let us subtract the second of these resulting equations from the first and integrate the difference over the interval $(0, l)$. We then obtain

$$\int_0^l [(pX')' Y_\varepsilon - (pY_\varepsilon')' X]\, dx$$
$$= \int_0^l [Y_\varepsilon(x, x_0) f(x) - X(x) g_\varepsilon(x, x_0)]\, dx.$$

Since the functions $X(x)$ and $Y_\varepsilon(x, x_0)$ vanish at the two limits of integration, the left member of this equation is equal to zero, which is easily seen by twice integrating by parts:

* In equation (24.3), the right side has two points of discontinuity of the first kind, namely, at $x = x_0 \pm \epsilon/2$. It can be shown that if $q \geqslant 0$ there exists a unique solution of equation (24.3) that satisfies the boundary conditions (24.2) and that is continuous and possesses a continuous first derivative on the interval $0 \leqslant x \leqslant l$. The second derivative has a discontinuity of the first kind at $x = x_0 \pm \epsilon/2$.

$$\int_0^l (pX')'\, Y_\varepsilon\, dx = pX'Y_\varepsilon \Big|_0^l - \int_0^l pX'Y'_\varepsilon\, dx$$

$$= -pY'_\varepsilon X \Big|_0^l + \int_0^l (pY'_\varepsilon)'\, X\, dx = \int_0^l (pY'_\varepsilon)'\, X\, dx.$$

Consequently,

$$\int_0^l Y_\varepsilon (x,\ x_0) f(x)\, dx = \int_0^l g_\varepsilon (x,\ x_0)\, X(x)\, dx$$

$$= \frac{1}{\varepsilon} \int_{x_0 - \frac{\varepsilon}{2}}^{x_0 + \frac{\varepsilon}{2}} X(x)\, dx \approx X(x_0). \quad (24.4)$$

If we assume that the function $Y_\varepsilon (x,\ x_0)$ approaches some limiting function (which we denote by $G(x,\ x_0)$) uniformly with respect to x as ε approaches 0, then, by taking the limit of the first and last members of (24.4), we obtain

$$X(x_0) = \int_0^l G(x,\ x_0) f(x)\, dx. \qquad (24.5)$$

This limiting function $G(x,\ x_0)$ is called Green's function for equation (24.1).

This nonrigorous reasoning does not yet enable us to give an exact proof of any facts at all. Therefore, we shall define Green's function independently of the comments that we have just made, and we shall show, first, that such a function exists and, second, that formula (24.5) is valid.

Before giving a precise definition of Green's function, let us show clearly what properties the limit $Y_\varepsilon (x, x_0)$ must possess if such a limit exists. Let us replace Y with $Y_\varepsilon (x,\ x_0)$ in the identity (24.3), and then, let us integrate this identity with respect to x from $x_0 - \delta$ to $x_0 + \delta$, where $\delta > \frac{\varepsilon}{2}$. This will give us

$$\int_{x_0-\delta}^{x_0+\delta} \{[pY'_\varepsilon (x,\ x_0)]' - qY_\varepsilon (x,\ x_0)\}\, dx = \int_{x_0-\delta}^{x_0+\delta} g_\varepsilon (x,\ x_0)\, dx = 1.$$

The first term can be integrated in explicit form, so

that we have

$$pY'_\varepsilon(x, x_0)\Big|_{x_0-\delta}^{x_0+\delta} - \int_{x_0-\delta}^{x_0+\delta} qY_\varepsilon(x, x_0)\,dx = 1.$$

If we now assume the validity of formally passing to the limit as ε approaches 0 with δ fixed, we obtain the equation

$$p(x_0+\delta)\,G'_x(x_0+\delta, x_0) - p(x_0-\delta)\,G'_x(x_0-\delta, x_0) - \int_{x_0-\delta}^{x_0+\delta} q(x)\,G(x, x_0)\,dx = 1,$$

which is valid for every $\delta > 0$. If we now take the limit as δ approaches 0 and assume that $p(x)$, $q(x)$, and $G(x, x_0)$ are continuous functions, we obtain the equation

$$p(x_0)\,[G'_x(x_0+0, x_0) - G'_x(x_0-0, x_0)] = 1.$$

From this it is clear that under the assumptions made, the derivative $G'_x(x, x_0)$ of Green's function with respect to x must have a saltus equal to $\frac{1}{p(x_0)}$ at $x = x_0$.

2. We shall now give a formal definition of Green's function for equation (24.1) and we shall prove that such a function exists.

The Green's function for equation (24.1) with boundary conditions (24.2) is the function $G(x, s)$ defined on the square $0 \leqslant x \leqslant l$, $0 \leqslant s \leqslant l$ that satisfies the following conditions:

(1) $G(x, s)$ as a function of x for $x \neq s$ is continuous and has continuous first and second derivatives and it satisfies the homogeneous equation

$$[pG'_x(x, s)]'_x - qG(x, s) = 0. \tag{24.6}$$

(2) $$G(0, s) = G(l, s) = 0.$$

(3) $G(x, s)$ is continuous in the square $0 \leqslant x \leqslant l, 0 \leqslant s \leqslant l$, and $G'_x(x, s)$ as a function of x has a discontinuity of the first kind with a saltus $\frac{1}{p(s)}$ at $x = s$; that is

$$G'_x(s+0,\ s)-G'_x(s-0,\ s)=\frac{1}{p(s)}\qquad (0<s<l).$$

In proving the existence of such a function, we assume that $q\geqslant 0$ so that $\lambda=0$ is an eigenvalue of the equation

$$(pX')'-qX+\lambda\rho X=0$$

with the boundary conditions (24.2). [See Section 39 of my Lektsii po teorii obyknovennykh differentsial'nykhuravnenii (Lectures on the theory of ordinary differential equations) 1952.]

With this assumption, the existence of the Green's function is simply shown by its construction. Specifically, suppose that $X_1(x)$ is any nontrivial solution of the equation $(pX')'-qX=0$, satisfying the condition

$$X_1(0)=0,$$

and that $X_2(x)$ is a nontrivial solution of the same equation satisfying the condition

$$X_2(l)=0.$$

Because of the assumption made, the solutions $X_1(x)$ and $X_2(x)$ are linearly independent. If this were not the case, they would be proportional to each other and each of them would vanish at $x=0$ and $x=l$ without being identically equal to zero, which is impossible since $\lambda=0$ is not an eigenvalue. Let us set

$$G(x,\ s)=\begin{cases} A(s)\,X_1(x), & 0\leqslant x\leqslant s,\\ B(s)\,X_2(x), & s<x\leqslant l.\end{cases}\tag{24.7}$$

Then, conditions (1) and (2) are satisfied for an arbitrary choice of $A(s)$ and $B(s)$.

Let us now choose $A(s)$ and $B(s)$ so that condition (3) will be satisfied. Since $G(x,\ s)$ is continuous at $x=s$ we have

$$A(s)\,X_1(s)=B(s)\,X_2(s),$$

from which we get

$$A(s)=c(s)\,X_2(s),$$
$$B(s)=c(s)\,X_1(s).$$

We require that the saltus in the derivative at the point

$x=s$ should have the given value $\frac{1}{p(s)}$:

$$G'_x(s-0,\ s)=c(s)X_2(s)X'_1(s),$$
$$G'_x(s+0,\ s)=c(s)X_1(s)X'_2(s),$$

from which we obtain

$$c(s)=\frac{1}{p(s)[X_1(s)X'_2(s)-X'_1(s)X_2(s)]}.$$

The denominator $p(s)[X_1(s)X'_2(s)-X_2(s)X'_1(s)]$ is independent of s. This is true because the expression in the square brackets is the Wronskian determinant $\Delta(X_1,\ X_2)$ of the linearly independent solutions $X_1(s)$ and $X_2(s)$. According to a well-known formula,

$$\Delta(X_1,\ X_2)=\Delta_0 e^{-\int_0^x \frac{p'(x)\,dx}{p(x)}}=\frac{\Delta_0 p(0)}{p(x)},$$

from which we see that $c(s)$ is constant.

Therefore, Green's function has the form

$$\left.\begin{aligned} G(x,\ s)&=\frac{1}{\Delta_0 p(0)}X_2(s)X_1(x) \qquad \text{for } 0\leqslant x\leqslant s,\\ G(x,\ s)&=\frac{1}{\Delta_0 p(0)}X_1(s)X_2(x) \qquad \text{for } s\leqslant x\leqslant l.\end{aligned}\right\} \tag{24.8}$$

This proves the existence of a Green's function.

It is immediately obvious from formula (24.8) that the Green's function is symmetric with respect to its arguments:

$$G(x,\ s)=G(s,\ x).$$

Let us now prove formula (24.5) for the solution $X(x)$ of equation (24.1) that satisfies the boundary conditions (24.2). Let us first show that the function

$$X(x)=\int_0^l G(x,\ s)f(s)\,ds \tag{24.9}$$

satisfies equation (24.1). Because of the symmetry of Green's function, the function defined by formula (24.9) coincides with (24.5). To evaluate $X'(x)$, let us represent (24.9) in the form

$$X(x)=\int_0^x G(x, s) f(s)\, ds+\int_x^l G(x, s) f(s)\, ds. \tag{24.10}$$

If we differentiate this equation with respect to x, we obtain

$$X'(x)=\int_0^x G_x'(x, s) f(s)\, ds+\int_x^l G_x'(x, s) f(s)\, ds$$
$$+G(x, x-0) f(x)-G(x, x+0) f(x).$$

Because of the continuity of Green's function, we have

$$X'(x)=\int_0^x G_x'(x, s) f(s)\, ds+\int_x^l G_x'(x, s) f(s)\, ds. \tag{24.11}$$

Differentiating this with respect to x, we obtain an expression for $X''(x)$ in the form

$$X''(x)=\int_0^x G_{xx}''(x, s) f(s)\, ds+\int_x^l G_{xx}''(x, s) f(s)\, ds$$
$$+G_x'(x, x-0) f(x)-G_x''(x, x+0) f(x).$$

Since

$$G_x'(x+0, x)-G_x'(x-0, x)=\frac{1}{p(x)},$$

we have

$$G_x'(x, x-0)-G_x'(x, x+0)=\frac{1}{p(x)}.$$

Therefore,

$$X''(x)=\int_0^l G_{xx}''(x, s) f(s)\, ds+\frac{f(x)}{p(x)}. \tag{24.12}$$

Substituting into equation (24.1) the expressions for X, X', and X'', we obtain

$$(pX')'-qX=\int_0^l [(p(x)\, G_x')_x'-qG] f(s)\, ds+f(x)=f(x).$$

It is clear from the form of the right member of equation (24.5) that the function $X(x)$ defined by equation (24.5) vanishes at $x=0$ and $x=l$.

Thus, formula (24.5) gives a solution of equation (24.1) satisfying conditions (24.2). Because of the assumption that $q \geqslant 0$, this solution of equation (24.1) is unique.

3. Let us show how Green's function for equation (24.1) can be used to reduce the eigenvalue problem examined in the preceding sections to an integral equation. To do this, we write the fundamental equation (22.1) in the form

$$(pX')' - qX = -\lambda\rho X \tag{24.13}$$

and, by setting $f(x) = -\lambda\rho X$, we apply formula (24.5) to it. We then obtain the equation

$$X(s) + \lambda \int_0^l G(x, s)\, \rho(x)\, X(x)\, dx = 0, \tag{24.14}$$

which is Fredholm's homogeneous equation of the second kind with symmetric kernel and parameter λ.

The kernel of equation (24.14) can be symmetrised by multiplying that equation by $\sqrt{\rho(s)}$. Then, this equation becomes an equation with an unknown function $\sqrt{\rho(s)}\,X(s)$ and symmetric kernel $G(x, s)\sqrt{\rho(x)\,\rho(s)}$. On the basis of formula (24.5), equation (24.13) with the boundary conditions $X(0) = X(l) = 0$ and equation (24.14) are equivalent in the sense that every solution of (24.13) that vanishes at $x=0$ and $x=l$ is a solution of equation (24.14) with the same λ and vice versa.

On the other hand, for equations of the form (24.14), the theorems on the existence of eigenvalues and eigenfunctions and the orthogonality of the system of eigenfunctions that were proven in Section 22 remain valid as does the theorem on the possibility of series expansion. [See, for example, my Lektsii po teorii integral'nykh uravnenii (Lectures on the theory of integral equations), Gostekhizdat, 1951, Sections 11-14.] The theorems on the existence and orthogonality of the eigenfunctions and the expansion theorem, which were proven in Section 22, follow immediately. True, to prove

that the function $f(x)$ can be expanded, we need to require continuity of its second derivative in order to represent it in the form (24.5) and to use the Hilbert-Schmidt theorem.

A Green's function which reduces the solution of a differential equation to that of an integral equation can also be defined for other types of boundary conditions and even for equations with several independent variables. However, as a rule, we can get an effective expression of it only for very special types of equations and boundary conditions.

4. With the use of Green's function, we can give the basis of the Fourier method of solving the mixed problem for equation (21.1) under conditions (23.1)-(23.2) without using the results of Section 22.

For simplicity, let us consider an equation of the type (23.19), where $p > 0$ and $q \geqslant 0$ and let us prove the theorem stated in subsection 1 of Section 23 for this equation. The series (21.12) in this case takes the form

$$\sum_{k=1}^{\infty} X_k(x)\left(A_k \cos\sqrt{\lambda_k}t + \frac{B_k}{\sqrt{\lambda_k}}\sin\sqrt{\lambda_k}t\right). \tag{24.15}$$

Here, the λ_k are eigenvalues and the $X_k(x)$ are the eigenfunctions of the equation

$$L(X) \equiv (pX')' - qX = -\lambda X \tag{24.16}$$

with boundary conditions (24.2). The existence of eigenvalues and eigenfunctions follows from the fact that equation (24.16) with conditions (24.2) is equivalent to the integral equation with symmetric kernel

$$X(x) + \lambda\int_0^l G(x, s)X(s)\,ds = 0, \tag{24.17}$$

where $G(x, s)$ is Green's function for the problem (24.16)-(24.2).

Since the initial functions $\varphi_0(x)$ and $\varphi_1(x)$ satisfy the conditions of the theorem in subsection 1 of Section 23, the conditions (see Section 23)

$$A_k = \int_0^l \varphi_0 X_k\,dx = -\frac{1}{\lambda_k}\int_0^l L(\varphi_0)X_k\,dx; \tag{24.18}$$

$$B_k = \int_0^l \varphi_1 X_k\, dx = -\frac{1}{\lambda_k} \int_0^l L(\varphi_1) X_k\, dx. \tag{24.19}$$

hold for the coefficients A_k and B_k in (24.15).

To show that the series (24.15) and the series that are obtained by twice differentiating termwise with respect to x and t converge uniformly, it will be sufficient to show that the following series converge uniformly on the interval $0 \leqslant x \leqslant l$

$$\sum_{k=1}^{\infty} |X_k(x)| (\lambda_k |A_k| + \sqrt{\lambda_k} |B_k|), \tag{24.20}$$

$$\sum_{k=1}^{\infty} |X_k'(x)| \left(|A_k| + \frac{|B_k|}{\sqrt{\lambda_k}} \right), \tag{24.21}$$

$$\sum_{k=1}^{\infty} |X_k''(x)| \left(|A_k| + \frac{|B_k|}{\sqrt{\lambda_k}} \right). \tag{24.22}$$

From equation (24.16), we have

$$X_k'' = \frac{-p'}{p} X_k' + \frac{q-\lambda_k}{p} X_k;$$

Consequently, the uniform convergence of the series (24.22) will follow from the uniform convergence of the series (24.20) and (24.21).

As before, let

$$D(f, g) = \int_0^l (pf'g' + qfg)\, dx, \qquad D(f) = \int_0^l (pf'^2 + qf^2)\, dx.$$

Obviously, $D(f) \geqslant 0$ for an arbitrary function f. If we multiply both sides of equation (24.16) by X_k and integrate from 0 to l, we obtain, upon integrating by parts,

$$\lambda_k \int_0^l X_k^2\, dx = D(X_k),$$

Since $X_k'(x) \not\equiv 0$, it follows that $\lambda_k > 0$ (for $k = 1, 2, \ldots$).

Lemma. Suppose that the function $f(x)$ is continuous on the interval $[0, l]$, that it vanishes at $x = 0$ and $x = l$,

and that it has a piecewise-continuous square-integrable derivative on that interval. Then,

$$\sum_{k=1}^{\infty} \lambda_k c_k^2 \leqslant D(f), \tag{24.23}$$

where $c_k = \int_0^l f X_k \, dx$ (for $k = 1, 2, \ldots$).

Proof: Integrating by parts, we obtain

$$D(f, X_k) = -\int_0^l f[(pX_k')' - qX_k]\,dx = \lambda_k c_k;$$

$$D(X_i, X_k) = \lambda_k \int_0^l X_i X_k \, dx = \lambda_k \delta_{ik}.$$

From this, we obtain

$$0 \leqslant D\left(f - \sum_{k=1}^{N} c_k X_k\right) = D(f) + D\left(\sum_{k=1}^{N} c_k X_k\right) - 2D\left(f, \sum_{k=1}^{N} c_k X_k\right) = D(f) - \sum_{k=1}^{N} \lambda_k c_k^2,$$

from which (24.23) follows.

By assumption, the function $L(\varphi_0)$ satisfies the conditions of the lemma just proven. Therefore inequality (24.23) holds for $L(\varphi_0)$. By using (24.18), we obtain

$$\sum_{k=1}^{\infty} \lambda_k^3 A_k^2 \leqslant D(L(\varphi_0)). \tag{24.24}$$

The function $L(\varphi_1)$ is continuous on the interval $[0, l]$. From the relation (24.19) and from Bessel's inequality, we obtain

$$\sum_{k=1}^{\infty} \lambda_k^2 B_k^2 \leqslant \int_0^l [L(\varphi_1)]^2 dx. \tag{24.25}$$

From equation (24.17), we have

$$\frac{X_k(x)}{\lambda_k} = -\int_0^l G(x, s) X_k(s)\, ds. \tag{24.26}$$

Consequently, for fixed x, the ratio $\frac{X_k(x)}{\lambda_k}$ is the kth Fourier coefficient of the function $G(x, s)$, which satisfies the conditions of the lemma on the interval $0 \leqslant s \leqslant l$. Therefore,

$$\sum_{k=1}^{\infty} \frac{X_k^2(x)}{\lambda_k} \leqslant D(G) \leqslant M_1 \quad \text{for} \quad 0 \leqslant x \leqslant l. \tag{24.27}$$

When we differentiate (24.26), we obtain

$$\frac{X_k'(x)}{\lambda_k} = -\int_0^l G_x X_k(s)\, ds.$$

It follows from this on the basis of Bessel's inequality that

$$\sum_{k=1}^{\infty} \frac{X_k'^2(x)}{\lambda_k^2} \leqslant \int_0^l G_x'^2\, ds \leqslant M_2 \quad \text{for} \quad 0 \leqslant x \leqslant l. \tag{24.28}$$

Let us now show that the series (24.20) converges uniformly on the interval $0 \leqslant x \leqslant l$. By applying Cauchy's inequality and using the estimate (24.27) we obtain

$$\sum_{k=n}^{n+m} |X_k(x)| (\lambda_k |A_k| + \sqrt{\lambda_k} |B_k|)$$

$$= \sum_{k=n}^{n+m} \frac{|X_k(x)|}{\sqrt{\lambda_k}} (\lambda_k^{\frac{3}{2}} |A_k| + \lambda_k |B_k|)$$

$$\leqslant \sqrt{\sum_{k=n}^{n+m} \frac{X_k^2(x)}{\lambda_k}} \left(\sqrt{\sum_{k=n}^{n+m} \lambda_k^3 A_k^2} + \sqrt{\sum_{k=n}^{n+m} \lambda_k^2 B_k^2} \right)$$

$$\leqslant \sqrt{M_1} \left(\sqrt{\sum_{k=n}^{n+m} \lambda_k^3 A_k^2} + \sqrt{\sum_{k=n}^{n+m} \lambda_k^2 B_k^2} \right).$$

From this it follows that the series (24.20) converges uniformly, so that the numerical series $\sum_{k=1}^{\infty} \lambda_k^3 A_k^2$ and $\sum_{k=1}^{\infty} \lambda_k^2 B_k^2$ converge by virtue of (24.24) and (24.25).

Let us prove that the series (24.21) converges uniformly. Because of (24.28), we have

$$\sum_{k=n}^{n+m} |X_k'(x)| \left(|A_k| + \frac{|B_k|}{\sqrt{\lambda_k}} \right)$$

$$= \sum_{k=n}^{n+m} \frac{|X_k'(x)|}{\lambda_k} (\lambda_k |A_k| + \sqrt{\lambda_k} |B_k|)$$

$$\leqslant \sqrt{M_2} \left(\sqrt{\sum_{k=n}^{n+m} \lambda_k^2 A_k^2} + \sqrt{\sum_{k=n}^{n+m} \lambda_k B_k^2} \right). \qquad (24.29)$$

The series $\sum_{k=1}^{\infty} \lambda_k^2 A_k^2$ converges by virtue of Bessel's inequality for the function $L(\varphi_0)$. The series $\sum_{k=1}^{\infty} \lambda_k B_k^2$ converges by virtue of the lemma (applied to the function φ_1). Therefore, it follows from (24.29) that the series (24.21) converges uniformly.

Remark 1: By following the same line of reasoning, we can show the basis of the Fourier method for solving the mixed problem for the general equation (21.1) if we use an estimate of the form (23.7) for the functions $T_k^*(t)$ and $T_k^{**}(t)$ and their derivatives.

2: The results of the present subsection yield the fundamental theorem of subsection 10 of Section 22 by following a different path. Indeed, for a function satisfying the conditions of the lemma of the present subsection, we obtain on the basis of inequalities (24.23) and (24.27)

$$\sum_{k=n}^{n+m} |c_k| |X_k(x)| = \sum_{k=n}^{n+m} \sqrt{\lambda_k} |c_k| \frac{|X_k(x)|}{\sqrt{\lambda_k}}$$

$$\leqslant \sqrt{D(G)} \sqrt{\sum_{k=n}^{n+m} \lambda_k c_k^2} < \varepsilon$$

for $n > N(\varepsilon)$. In other words, the series $\sum_{k=1}^{\infty} c_k X_k(x)$ converges absolutely and uniformly on the interval $[0, l]$.

25. STUDY OF THE VIBRATIONS OF A MEMBRANE

1. In Section 1, we considered as an example the equation of the vibrations of a membrane

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. \tag{25.1}$$

Suppose that in its position of equilibrium, the membrane coincides with some bounded region G of the xy-plane that has a piecewise smooth boundary Γ. Then the function $u(t, x, y)$ representing these vibrations must satisfy equation (25.1) and the initial conditions

$$\begin{aligned} u(0, x, y) &= \varphi_0(x, y) \text{ (initial deviation)}, \\ u'_t(0, x, y) &= \varphi_1(x, y) \text{ (initial velocity)}, \end{aligned} \tag{25.2}$$

when the point $(x, y) \in G$. Also, on the boundary Γ of the region G, the function $u(t, x, y)$ must satisfy certain boundary conditions of the type considered in Section 1.

Let us consider the simplest case, that of a membrane that is firmly fastened at its edge. Then, the boundary condition is

$$u(t, x, y) = 0, \text{ when } (x, y) \in \Gamma. \tag{25.3}$$

If we solve this problem (again by the method of separating the variables), we obtain

$$u(t, x, y) = T(t)\, v(x, y).$$

By a procedure analogous to that of the one-dimensional case, we obtain the following equations for the functions $T(t)$ and $v(x, y)$:

$$T'' + \lambda T = 0, \tag{25.4}$$

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \lambda v = 0. \tag{25.5}$$

For equation (25.5) with boundary condition (25.3), there exists an infinite sequence of eigenvalues. The eigenfunctions corresponding to the different eigenvalues are orthogonal to each other. In contrast with the case of a single independent variable, to certain eigenvalues there may correspond not one but several linearly independent eigenfunctions. Such eigenvalues are said to be multiple-valued. From among the eigenfunctions corresponding to a given eigenvalue, we can always choose a finite system of linearly independent and mutually orthogonal eigenfunctions such that each eigen-

function belonging to the eigenvalue in question can be expressed as a linear combination of the others.

The set of eigenfunctions chosen in this way corresponding to all the eigenvalues forms a complete orthogonal system of functions

$$v_1(x, y), \quad v_2(x, y), \quad \dots, \quad v_n(x, y), \quad \dots$$

Let us expand the functions $\varphi_0(x, y)$ and $\varphi_1(x, y)$ in series of functions $v_n(x, y)$:

$$\varphi_0(x, y) = \sum_{n=1}^{\infty} A_n v_n(x, y), \quad \varphi_1(x, y) = \sum_{n=1}^{\infty} B_n v_n(x, y) \tag{25.6}$$

Let us choose two linearly independent solutions $T^*(t)$ and $T^{**}(t)$ of equation (25.4) that will satisfy the conditions

$$T^*(0) = 1; \; T^{*\prime}(0) = 0; \; T^{**}(0) = 0; \; T^{**\prime}(0) = 1.$$

The series

$$u(t, x, y) = \sum_{n=1}^{\infty} v_n(x, y)\,[A_n T_n^*(t) + B_n T_n^{**}(t)] \tag{25.7}$$

represents the solution of our problem in the case in which this series and the series that are obtained from it by twice differentiating termwise with respect to t, x and y converge uniformly.

We shall stop here for two particular cases in which the eigenfunctions of equation (25.5) can in their turn be found by the method of separation of variables. One may proceed in an analogous manner in the case of a larger number of independent variables. These cases can be completely investigated by reducing them to the one-dimensional eigenvalue problem by means of the following lemma.

Lemma. Suppose that $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ constitute a complete system of orthogonal functions on the interval $[a, b]$ that are normalised with weight $\rho_1(x)$. Suppose also that for each n $(n = 1, 2, \dots)$, there is a complete system of orthogonal functions

$$\psi_{n1}(y), \; \psi_{n2}(y), \; \dots, \; \psi_{nm}(y), \dots \tag{25.8}$$

on the interval $[c, d]$ that are normalised with weight $\rho_2(y)$ The functions $\rho_1(x)$ and $\rho_2(x)$ are assumed to be

continuous and non-negative. In this case, the functions

$$X_{nm}(x, y) = \varphi_n(x)\psi_{nm}(y)$$

constitute a complete system of orthogonal functions in the rectangle $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$ that are normalised with weight $\rho(x, y) = \rho_1(x)\rho_2(y)$. In other words, the equations

$$\int_c^d \int_a^b \rho(x, y) X_{nm}(x, y) X_{n'm'}(x, y)\, dx\, dy = \begin{cases} 1 & \text{for } n = n', \quad m = m', \\ 0 & \text{for } n \neq n' \quad \text{or} \quad m \neq m', \end{cases} \tag{25.9}$$

hold, and if

$$c_{nm} = \int_a^d \int_c^b \rho(x, y) f(x, y) X_{nm}(x, y)\, dx\, dy,$$

then, for an arbitrary function $f(x, y)$ that is continuous in the rectangle referred to above, Parseval's equation

$$\int_c^d \int_a^b \rho(x, y)[f(x, y)]^2\, dx\, dy = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_{nm}^2 \tag{25.10}$$

holds.

Formula (25.9) may be proved valid by inspection. To prove (25.10), let us set

$$\int_a^b \rho_1(x) f(x, y) \varphi_n(x)\, dx = g_n(y).$$

Then, it is obvious that

$$\int_c^d \rho_2(y) g_n(y) \psi_{nm}(y)\, dy = c_{nm},$$

$$\int_a^b \rho_1 [f(x, y)]^2\, dx = \sum_{n=1}^{\infty} g_n^2(y)$$

and that

$$\int_c^d \rho_2(y) g_n^2(y)\, dy = \sum_{m=1}^{\infty} c_{nm}^2,$$

since the systems $\psi_{nm}(y)$ are complete for arbitrary n and since the function $g_n(y)$ is square-integrable.

Since the series

$$\sum_{n=1}^{\infty} g_n^2(y)$$

consists of positive terms and converges to a continuous function at every point of the interval $[c, d]$ it follows from Dini's theorem that it converges uniformly on that interval. Therefore,

$$\int_c^d \int_a^b \rho(x, y)[f(x, y)]^2 \, dx \, dy = \int_c^d \rho_2(y) \sum_{n=1}^{\infty} g_n^2(y) \, dy$$

$$= \sum_{n=1}^{\infty} \int_c^d \rho_2(y) g_n^2(y) \, dy = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} c_{nm}^2,$$

which completes the proof.

2. Let us now consider the first particular case referred to above, namely, the vibrations of a rectangular membrane.

Suppose that the region is a rectangle $0 \leqslant x \leqslant a, \ 0 \leqslant y \leqslant b$. Separating the variables in equation (25.5), we obtain

$$v(x, y) = X(x) \, Y(y).$$

When we substitute these functions into equation (25.5), it will assume the form

$$X''Y + XY'' + \lambda XY = 0.$$

When we divide by XY and transpose $\frac{X''}{X}$ to the left side of the equation, we obtain

$$\frac{Y''}{Y} + \lambda = -\frac{X''}{X}$$

which is equivalent to the two ordinary differential equations

$$X'' + \alpha X = 0, \quad Y'' + \beta Y = 0,$$

where α is a constant and $\beta = \lambda - \alpha$. In accordance with the boundary condition (25.3), the first of these equations must be solved under the conditions

$$X(0) = X(a) = 0,$$

and the second under the analogous conditions

$$Y(0) = Y(b) = 0.$$

To satisfy these conditions, we must assume that α and β are both positive (see Section 20). By repeating the reasoning of Section 20, we obtain sequences of eigenvalues and normalised eigenfunctions of the two equations above:

$$\alpha_n = \frac{n^2\pi^2}{a^2}, \quad X_n(x) = \sqrt{\frac{2}{a}} \sin \frac{n\pi}{a} x;$$

$$\beta_m = \frac{m^2\pi^2}{b^2}, \quad Y_m(y) = \sqrt{\frac{2}{b}} \sin \frac{m\pi}{b} y;$$

$$(n, \ m = 1, \ 2, \ldots).$$

According to the lemma, the system of functions

$$v_{nm}(x, \ y) = \frac{2}{\sqrt{ab}} \sin \frac{n\pi}{a} x \sin \frac{m\pi}{b} y \tag{25.11}$$

is a complete system of orthogonal normalised solutions of equation (25.5) for the rectangle $0 \leqslant x \leqslant a$, $0 \leqslant y \leqslant b$ with boundary conditions (25.3). (Here, $Y_{nm}(y) = Y_m(y)$ for arbitrary n.) To each function $v_{nm}(x, y)$ there corresponds the eigenvalue

$$\lambda_{nm} = \pi^2 \left(\frac{n^2}{a^2} + \frac{m^2}{b^2} \right).$$

Obviously, if the numbers a and b are commensurable, we can get the same λ for different choices of n and m, that is, for different eigenfunctions. Thus, we have here an example of multiple eigenvalues.

The question of expanding the initial conditions in a series of functions of the form (25.11) is the well-studied question of expanding functions in a double Fourier sine series. Suppose that the initial conditions are extended as odd functions of x and y to the rectangle $|x| \leqslant a$, $|y| \leqslant b$ and as periodic functions over the entire plane. Suppose that they are then four times continuously differentiable functions. Then, the coefficients in the expansions (25.6) converge to zero with sufficient speed that the series (25.7) can be twice differentiated. Thus, in this case the Fourier method for solving the given problem is completely justified. We

see that an arbitrary vibration of a membrane, just as the vibration of a string, can be represented as a series expansion of simple, so-called 'natural' vibrations corresponding to the eigenvalues λ_{nm}.

Of interest are the 'nodal curves' of such vibrations. These are the curves along which the eigenfunction corresponding to a given eigenvalue vanishes. Let us consider these curves in the case of a rectangular membrane. If the given eigenvalue is not multiple, that is, if only one eigenfunction

$$\frac{2}{\sqrt{ab}} \sin \frac{n\pi}{a} x \sin \frac{m\pi}{b} y,$$

corresponds to it, then the nodal lines will be simply segments of straight lines parallel to the sides of the rectangle. On the other hand, if the eigenvalue is multiple, different nodal lines will correspond to the different combinations of the eigenfunctions belonging to it and their form may be quite varied. Figure 11 shows the nodal curves of a square membrane whose side is of unit length that correspond to the values $\lambda = 5\pi^2$, $10\pi^2$, $13\pi^2$, and $17\pi^2$. Under the drawings of the nodal curves, are shown the corresponding eigenfunctions.

3. As our second example, let us consider a circular membrane. To study it, it is natural to write equation (25.5) in polar coordinates.

Setting $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, we obtain

$$\frac{\partial^2 v}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial v}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 v}{\partial \varphi^2} + \lambda v = 0. \tag{25.12}$$

If we place the coordinate origin at the centre of the circle D representing the edge (fixed) of the membrane and if for simplicity we take the radius of this circle as being equal to unity, the boundary condition (25.3) takes the form

$$v(1, \varphi) = 0.$$

Applying the method of separation of variables, we set

$$v(\rho, \varphi) = R(\rho)\Phi(\varphi).$$

When we substitute this expression into (25.12) and separate

the variables, we obtain the following two ordinary differential equations for $R(\rho)$ and $\Phi(\varphi)$:

$$\Phi''(\varphi) + \alpha\Phi(\varphi) = 0, \tag{25.13}$$

$$\rho^2 R''(\rho) + \rho R'(\rho) + (\lambda\rho^2 - \alpha) R = 0. \tag{25.14}$$

In solving equation (25.13) we get the condition of periodicity form the physical properties of the problem. We are

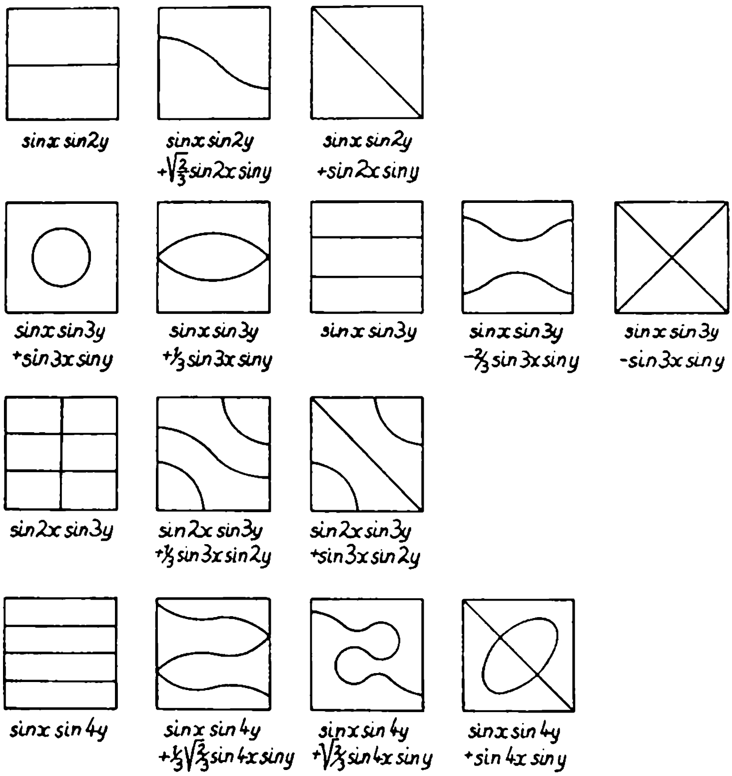


Fig. 11

interested only in solutions with the period 2 . Such solutions exist for

$$\alpha = 0,\ 1^2,\ 2^2, \dots,\ n^2, \dots$$

For these values of α,

$$\Phi_n(\varphi) = A_n \cos n\varphi + B_n \sin n\varphi.$$

We can choose a complete system of orthogonal functions $\Phi_n(\varphi)$ that are normalised on the circle, for example,

$$\Phi_0^* = \frac{1}{\sqrt{2\pi}}; \quad \Phi_n^*(\varphi) = \sqrt{\frac{1}{\pi}} \cos n\varphi; \quad \Phi_n^{**}(\varphi) = \sqrt{\frac{1}{\pi}} \sin n\varphi.$$

Let us return to equation (25.14). After making the substitution $\alpha = n^2$ and the change of independent variable

$$\rho_1 = \rho\sqrt{\lambda}$$

we obtain Bessel's equation of nth order

$$\rho_1^2 R''(\rho_1) + \rho_1 R'(\rho_1) + (\rho_1^2 - n^2) R(\rho_1) = 0.$$

Its only solution (up to a constant factor) that is bounded as $\rho_1 \to 0$ (that is, as $\rho \to 0$) is Bessel's function of nth order of the first kind†, denoted by $J_n(\rho_1)$.

As we know, for an arbitrary n, the function $J_n(x)$ has an infinite number of positive roots $\mu_1^{(n)}, \mu_2^{(n)}, \ldots \mu_m^{(n)}, \ldots$, so that $J_n(\mu_m^{(n)}) = 0$.

We also know that for an arbitrary fixed n, the functions $J_n(\mu_m^{(n)}x)$ (for $m = 1, 2, \ldots$) are mutually orthogonal with weight x on the interval $(0,1)$ and they form a complete system of orthogonal functions on that interval:

$$\int_0^1 x J_n(\mu_m^{(n)}x) J_n(\mu_{m_1}^{(n)}x)\, dx = 0 \quad \text{for} \quad m \neq m_1.$$

For arbitrary n, the functions

$$\varphi_{nm}(x) = \frac{J_n(\mu_m^{(n)}x)}{\sqrt{\int_0^1 x\,[J_n(\mu_m^{(n)}x)]^2\, dx}}$$

constitute a complete system of orthonormal functions. We shall not give proof of these facts‡ but only note that they are a generalisation of the properties of eigenfunctions shown in Section 22 to equations with more general coeffici-

† See, for example STEPANOV, V.V., *Course in differential equations*, Chapter VI, section 2, 250, Fizmatgiz (1959).

‡ See, for example, KUZ'MIN, R.O., *Bessel functions*. ONTI (1935) or TIKHONOV, A.N. and SAMARSKII, A.A., *The equations of mathematical physics*, Gostekhizdat, 566-619 (1953).

ents than was assumed in that section. Specifically, equation (25.14) can be written in the form

$$(\rho R')' - \frac{\alpha}{\rho} R + \lambda \rho R = 0,$$

and we see that the first and last coefficients vanish at the lower end point of the interval $[0, 1]$ and that the fraction $\frac{\alpha}{\rho}$ becomes infinite at that point. In connection with this, it can be shown that boundedness of the solution at $\rho = 0$ is a sufficient boundary condition of the eigenvalue problem for equation (25.14) for the solution to be determined up to a constant factor if at $\rho = 1$ any condition of the type (22.2) is given.

Let us require that

$$J_n(\sqrt{\lambda}\rho) = 0,$$

at $\rho = 1$, that is, we require that

$$J_n(\sqrt{\lambda}) = 0.$$

We see that, if $\mu_1^{(n)}, \mu_2^{(n)}, \ldots, \mu_m^{(n)}, \ldots$ is the sequence of zeros of the function $J_n(x)$, the eigenvalues λ of our problem will be

$$\lambda_{nm} = [\mu_m^{(n)}]^2,$$

and the normalised eigenfunctions of equation (25.14) will be the functions

$$\psi_{nm}(\rho) = J_n(\mu_m^{(n)}\rho) \frac{1}{\sqrt{\int\limits_0^1 \rho\,[J_n(\mu_m^{(n)}\rho)]^2\,d\rho}}.$$

Applying the lemma of subsection 1, we can obtain a complete system of eigenfunctions

$$\psi_{nm}(\rho)\,\Phi_n^*(\varphi), \quad \psi_{nm}(\rho)\,\Phi_n^{**}(\varphi)$$

of equation (25.12) and thus find the solution of our problem by expanding the functions $\varphi_0(\rho, \varphi)$ and $\varphi_1(\rho, \varphi)$ in series of the form

$$\varphi_0(\rho, \varphi) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} [c^*_{nm} \Phi^*_n(\varphi) + c^{**}_{nm} \Phi^{**}_n(\varphi)] \psi_{nm}(\rho),$$

$$\varphi_1(\rho, \varphi) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} [d^*_{nm} \Phi^*_n(\varphi) + d^{**}_{nm} \Phi^{**}_n(\varphi)] \psi_{nm}(\rho).$$

When we multiply the terms in the first sequence by the corresponding functions $T^*(t)$ and the terms in the second series by the corresponding function $T^{**}(t)$ and when we add the series that are thus obtained, we obtain the series (25.7), which represents the solution of the given problem. The uniform convergence and the possibility of termwise differentiation of this series will, as usual, be assured for sufficiently smooth functions $\varphi_0(\rho, \varphi)$ and $\varphi_1(\rho, \varphi)$ if they satisfy the same boundary conditions that the desired solution of equation (25.1) must satisfy plus certain additional conditions on the boundary of the circle.

26. ADDITIONAL REMARKS ON EIGENFUNCTIONS AND SOLVABILITY OF THE MIXED PROBLEM FOR HYPERBOLIC EQUATIONS

1. Everything that has been said up to now with regard to the eigenvalue problem for equation (22.1) can be carried over in a natural manner to the analogous problem for the equation

$$(p_1 u'_x)'_x + (p_2 u'_y)'_y - qu + \lambda \rho u = 0 \tag{26.1}$$

and the equation

$$(p_1 u'_x)'_x + (p_2 u'_y)'_y + (p_3 u'_z)'_z - qu + \lambda \rho u = 0 \tag{26.2}$$

under the assumption that the functions p_i, their derivatives, and ρ are continuous and that p_i and ρ are bounded below by some positive constant.

Let us seek solutions to equation (26.1) in some finite region G with a piecewise smooth boundary L that are not identically zero and that satisfy either the condition

$$u = 0 \tag{26.3}$$

or the condition on the boundary

$$\frac{\partial u}{\partial n} + \sigma u = 0. \tag{26.4}$$

We shall bear in mind, that in this case the partial differentiation operator $\frac{\partial}{\partial n}$ indicates the differentiation in the direction of the 'conormal'. This 'conormal' is defined at every point on the boundary as being the vector $[(p_1 \cos(n, x), p_2 \cos(n, y))]$, where $\cos(n, x)$, and $\cos(n, y)$ are the respective cosines of the angles which the direction of the outer normal makes with the x- and y-axes and where σ is some nonnegative function defined on the boundary of the finite region G. Analogously with the procedure which we already adopted above, let us determine first the eigenfunctions and the eigenvalues of this problem. We construct the functionals

$$H(u) = \iint_G \rho u^2 \, dx \, dy,$$

$$D(u) = \iint_G (p_1 u_x^2 + p_2 u_y^2 + q u^2) \, dx \, dy$$

in the case of the boundary condition (26.3) and

$$\begin{aligned} D^*(u) &= D(u) + \int_L \sigma u^2 \, dl \\ &= \iint_G (p_1 u_x^2 + p_2 u_y^2 + q u^2) \, dx \, dy + \int_L \sigma u^2 \, dl \end{aligned}$$

in the case of the boundary condition (26.4). Then, we can easily carry over to this case all the theorems regarding the properties of eigenfunctions and eigenvalues that were shown in Section 22.

In particular, Courant's theorem on the extreme-value property of the eigenfunctions and the consequent dependence of the eigenvalues on the coefficients of the equation, on the region G, and on the boundary conditions apply to this case. As can easily be seen, the nth eigenvalue does not decrease when the functions

$$\sigma(l), \; p_1, \; p_2, \; q, \; \frac{1}{\rho}$$

increase.

We arrive at problems of this type when, for example, we study the vibrations of a membrane.

Then, properties analogous to those described in subsections 5(c) and 6 of Section 22 have an interesting physical interpretation illustrating the nature of the change in frequency of the natural vibrations of the membrane when

it is fastened at certain portions of the region G (see subsection 5(c)) or when there is friction in the membrane (see subsection 6).

This last property corresponds to the familiar physical fact that objects emit a lower tone after they are broken than before.

For equations (26.1) and (26.2), the theorems on the completeness of the system of eigenfunctions and the possibility of expanding in absolutely and uniformly convergent series of eigenfunctions each function f that satisfies on the boundary the same boundary conditions as do the eigenfunctions that we have been considering also remain valid. However, in the theorem on the possibility of expansion, we now need to require greater smoothness of f than we did in the case of a single independent variable.

When there are two or three independent variables, it is sufficient to require that the function f have continuous first and second derivatives in a closed region and that the boundary of the region be sufficiently smooth. The method of proof of the expansion theorem that we used earlier for the case of a single indpendent variable is not applicable in these cases. Here, we need to use integral equations.

The eigenfunctions for general second-order elliptic equations have been investigated by M.V. Keldysh*. The properties of the eigenfunctions and the eigenvalues for such equations (the expansion theorem. the structure of the spectrum) are considerably more complicated than in the particular cases (22.1), (26.1). and (26.2) that we examined above.

2. When we were speaking of the behaviour of the eigenfunctions of equation (22.1), we did not treat the question of the frequency of change of sign (the ' zeros ') of the function $X_n(x)$ corresponding to the eigenvalue λ_n on the interval $(0, l)$. The so-called oscillation theorems of Sturm deal with this question.

It turns out that, first, the n th eigenfunction for equation (22.1) under the boundary conditions (22.5) has exactly $n-1$ zeros in the interval $[0, l]$ and that, second, the zeros of the function $X_{n+1}(x)$ alternate with the zeros of the function $X_n(x)$; that is, in every interval between two zeros of $X_{n+1}(x)$ lies one zero of the function $X_n(x)$ [cf. I.G.Petrovskii,

* KELDYSH, M.V., *Dokl. Akad. Nauk SSSR,* **77**, No. 1, 11-14 (1951).

Lektsii po teorii obyknovennykh differentsial'nykh uravnenii (Lectures on the theory of ordinary differential equations), Gostekhizdat, 1952, Section 39].

With regard to the nodal lines of the nth eigenfunction of equation (26.1) with the boundary conditions (26.3), it is shown that they partition the basic region G into no more than n subregions, and it is known that, in contrast with the case of a single independent variable there cannot be fewer than n of these regions. For the case of several independent variables, no theorems have been proven that are analogous to Sturm's theorem on the alternation of the zeros of successive eigenfunctions of an ordinary differential equation. We know even less about the asymptotic behaviour of the eigenfunctions for arbitrary regions.

3. Many problems in physics, both classical and modern, lead to a determination of the eigenfunctions and eigenvalues of the equation

$$u'' + \lambda u = R(x)\, u \tag{26.5}$$

in the interval $-\infty < x < \infty$ or in a finite interval $(0, l)$ but under the assumption that the function $R(x)$ becomes infinite at one or both end points of the interval*.

The theorem on the possibility of expansion in eigenfunctions is generalised in different ways according to the different cases. We shall present the two most important cases.

(a) A finite interval $0 < x < l$ with $R(0) = \infty$. In many problems, instead of a boundary condition at the point $x = 0$, we have the condition that

$$\int_0^l u^2(x, \lambda)\, dx < \infty, \tag{26.6}$$

where $u(x, \lambda)$ is a solution of equation (26.5). Here, it turns out that in certain cases not all solutions of the equation satisfy the condition (26.6). In many cases, condition (26.6) and the boundary condition at the point $x = l$ determines uniquely (up to a constant factor) the eigenvalues and eigen-

* We recall that it is possible to reduce an arbitrary equation of the form (22.1) with sufficiently smooth coefficients to this form by a change of variables.

functions. Furthermore, the spectrum is discrete and Sturm's oscillation theorem remains valid.

In other cases, it turns out that all solutions of equation (26.5) satisfy conditions (26.6). Then, condition (26.6) is insufficient for determining the spectrum of equation (26.5). A supplementary boundary condition at the point 0 is necessary. We shall not stop to explain the nature of this additional condition. Under this supplementary condition and the condition at $x=l$, the spectrum becomes discrete. In both cases, the theorem on the possibility of series expansion remains valid for a broad class of functions.

Both possibilities are clearly illustrated by Bessel's equation

$$y''+\frac{1}{x}y'+\left(s^2-\frac{\nu^2}{x^2}\right)y=0. \tag{26.7}$$

Its solutions are*

$$J_\nu(sx), \quad Y_\nu(sx)=\frac{J_\nu(sx)\cos\nu\pi-J_{-\nu}(sx)}{\sin\nu\pi}$$

If we make the substitution $y_1=\sqrt{x}\,y$, equation (26.7) becomes

$$y_1''+\left(s^2-\frac{\nu^2-\frac{1}{4}}{x^2}\right)y_1=0 \qquad (26.7)_1$$

which is of the type (26.5).

If $\nu\geqslant 1$, condition (26.6) is satisfied only by the function $\sqrt{x}\,J_\nu(sx)$. For $0\leqslant\nu<1$ all solutions of equation (26.7) satisfy this condition.

In the first case, to obtain the eigenfunctions and eigenvalues, we need to impose a boundary condition, for example the condition

$$\frac{d}{dx}\{\sqrt{x}\,J_\nu(sx)\}_{x=l}-H(\sqrt{x}\,J_\nu(sx))_{x=l}=0. \tag{26.8}$$

on the function $\sqrt{x}\,J_\nu(sx)$ only at the point l. This condition and condition (26.6) determine the eigenfunctions and eigenvalues.

In the second case, when $0\leqslant\nu<1$, in addition to the

* For ν an integer, the function $Y_\nu(sx)$ is defined as the limit of the expression the right as ν approaches the given integral value.

condition (26.8) we need to add some condition at the point $x=0$.

(b) The interval $(0, \infty)$ with $R(x)$ a continuous function. In this case, physical problems usually lead to a search for solutions $u(x)$ of equation (26.5) that satisfy some boundary condition at $x=0$ and that remain bounded as $x \to \infty$. In this case, if $R(x)$ is an absolutely integrable function in the interval $(0, \infty)$, we obtain what is called a continuous spectrum, that is, a continuous sequence of eigenvalues and a family of eigenfunctions $u(x, \lambda)$, that vary continuously with change in λ. Parseval's equation, that is, the definition of completeness of a system of eigenfunctions, is generalised to this case. We have the following

Theorem. Suppose that $f(x)$ is a square-integrable function on the interval $(0, \infty)$. Then,

$$\int_0^\infty f^2(x)\,dx = \int_{-\infty}^\infty F^2(\lambda)\,d\rho(\lambda) \quad \text{(Parseval's equation)}$$

where $F(\lambda)$ [the generalised Fourier transform of the function $f(x)$] is the limit of a sequence of functions

$$F_n(\lambda) = \int_0^n f(x)\,u(x, \lambda)\,dx,$$

that converge in means as $n \to \infty$

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} [F(\lambda) - F_n(\lambda)]^2 d\rho(\lambda) = 0.$$

Here, $\rho(\lambda)$ is some nondecreasing function.

The representiation of the function $f(x)$ in the form of an integral of eigenfunctions with respect to a function of the parameter λ, that is, a formula of the form

$$f(x) = \int_{-\infty}^{+\infty} u(x, \lambda)\,dg(\lambda)$$

for some function $g(\lambda)$ (the analogue of the ordinary Fourier integral for the equation $u'' + \lambda u = 0$), is valid under much stronger assumptions, which we shall not present here.

The reader can find a more detailed explanation of questions of this nature in the book by B.M. Levitan, Razlozhenie

po sobstvennym funktsiyam (Expansion in eigenfunctions), Gostekhizdat, 1950.

4. In the case of several dimensions, just as in that of a single independent variable, we sometimes need to examine the eigenvalue problem for an equation with coefficients that become infinite. There is no general theory at all regarding such problems but, in particular cases, mathematicians have succeeded in finding the complete solution to the problem and in exhibiting the expansion in the eigenfunctions of the problem. As an example, we cite the equation of the vibrations of a gas in space

$$\Delta u = u_{tt}.$$

In solving this equation by the Fourier method, we encounter the problem of determining the eigenfunctions of the equation

$$\Delta u + \lambda u = 0$$

for some region G. If the region G is a sphere of radius 1 with centre at the coordinate origin, then by converting the equation to spherical coordinates and finding solutions of the form

$$u(\rho, \theta, \varphi) = f(\rho)\, Y(\theta, \varphi),$$

we obtain for the function $Y(\theta, \varphi)$ the equation

$$\frac{1}{\sin\theta}\left[\frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\theta}\frac{\partial Y}{\partial\varphi}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial Y}{\partial\theta}\right)\right] + kY = 0,$$

the coefficients of which become infinite at the poles of the sphere $\theta = 0$ and $\theta = \pi$. For boundary conditions for this problem, we have the conditions that the solution be continuous and single-valued on the sphere $\rho = 1$. As in the case of continuous coefficients, we obtain under these conditions an infinite sequence of eigenvalues $k_n = n(n+1)$. To each eigenvalue k_n there correspond $2n+1$ linearly independent eigenfunctions $Y_n^m(\theta, \varphi)$ (the Legendre functions of nth order, $m = 1, 2, \ldots, 2n+1$). Here, the sequence of eigenvalues is complete on the surface of the sphere and every continuous sufficiently smooth function on the sphere can be expanded in a uniformly convergent series of Legendre functions.

*5. Variational methods for finding approximation of eigenfunctions and eigenvalues**. As was shown in Section 22, the problem of finding the first eigenvalue and the first eigenfunction of equation (22.1) with boundary conditions

$$X(0) = X(l) = 0$$

is equivalent to the problem of finding the minimum of the functional

$$D(X) = \int_0^l (pX'^2 + qX^2)\, dx \tag{26.9}$$

under the condition

$$H(X) = \int_0^l \rho X^2 dx = 1 \tag{26.10}$$

in the class of functions $X(x)$ that are continuously differentiable on the interval $[0, l]$ and that vanish at the end points of that interval. For an approximate solution of this problem, we use Ritz' method, which consists in the following: We consider an arbitrary system composed of an infinite number of linearly independent functions $\varphi_n(x)$, for $0 \leqslant x \leqslant l$, that are continuously differentiable and that satisfy the boundary conditions

$$\varphi(0) = \varphi(l) = 0.$$

We shall seek an approximate solution of this extreme-value problem in the form of a linear combination of a finite number of functions

$$X_N(x) = \sum_{n=1}^{N} a_n \varphi_n(x) \tag{26.11}$$

with undetermined coefficients a_n

When we substitute (26.11) into (26.9) and (26.10) and carry out the integration, we arrive at the problem of finding the minimum of the quadratic form

$$g(a_1, \ldots, a_N) = \sum_{n, m=1}^{N} a_m a_n \int_0^l [p\varphi_n'(x)\, \varphi_m'(x)$$

(continued on top of next page)

* See KANTOROVICH, L.V. and KRYLOV, V.I., *Approximate methods of higher analysis*, Gostekhizdat, Chapter IV, 258-373 (1952).

$$+ q\varphi_n(x)\varphi_m(x)]\,dx = \sum_{n,\,m=1}^{N} A_{mn}a_m a_n$$

under the condition

$$h(a_1, \ldots, a_N) = \sum_{n,\,m=1}^{N} a_m a_n \int_0^l \rho\varphi_n\varphi_m dx = \sum_{n,\,m=1}^{N} B_{mn}a_m a_n = 1.$$

This is now a problem in differential calculus, which in practice is easily solved since the derivatives of g and h with respect to a_k are linear functions of $a_1, \ldots, a_N$, so that the system of equations

$$\frac{\partial(g - \lambda h)}{\partial a_k} = 0 \qquad (k = 1,\ 2,\ \ldots,\ N) \tag{26.12}$$

is a linear homogeneous system of equations in a_k. The determinant of this system is an Nth-degree polynomial in λ. It vanishes at

$$\lambda = \lambda_1^{(N)}, \ldots, \lambda_N^{(N)}, \ \lambda_1^{(N)} \leqslant \lambda_2^{(N)} \leqslant \ldots \leqslant \lambda_N^{(N)}.$$

All the λ are real. For every $\lambda_i^{(N)}$, there exists a nontrivial solution $a_1^{(i)}$, $a_2^{(i)}$, $\ldots$, $a_N^{(i)}$ of the system (26.12). If $\lambda_i^{(N)}$ is a root of multiplicity k of the determinant, the system (26.12) has k linearly independent solutions $(a_1^{(i)}, \ldots, a_N^{(i)})$ for $\lambda = \lambda_i^{(N)}$.

Suppose that the system of functions $\varphi_n(x)$ is such that, for every continuously differentiable function $f(x)$ defined on the interval $[0, l]$ that satisfies the conditions $f(0) = f(l) = 0$ and for arbitrary $\varepsilon > 0$, it is possible to find a linear combination

$$\sum_{k=1}^{m} c_k\varphi_k$$

of functions φ_n with constant coefficients such that

$$|f(x) - \sum_{k=1}^{m} c_k\varphi_k(x)| < \varepsilon \text{ and } |f'(x) - \sum_{k=1}^{m} c_k\varphi_k'(x)| < \varepsilon$$

on the interval $[0, l]$. Then, for every fixed value of i,

$$\lambda_i^{(N)} \to \lambda_i, \text{ as } N \to \infty$$

where λ_i is the i th eigenvalue of the given problem. The numbers $\lambda_i^{(N)}$ for i not exceeding some fixed number M, and for N sufficiently large, are approximate values of the first M eigenvalues of the equation (22.1) under the boundary conditions

$$X(0) = X(l) = 0.$$

From the sequence of functions $X_i^{(N)}$ we can choose a subsequence $X_i^{(N')}$ such that

$$X_i^{(N')}(x) \longrightarrow X_i(x) \quad \text{when} \quad N' \longrightarrow \infty,$$

uniformly on the interval $[0, l]$ where $X_i(x)$ is the i th eigenfunction of the given problem.

The rapidity of covergence of $X_i^{(N')}(x)$ to $X_i(x)$ depends in a significant way* on the choice of the functions $\varphi_n(x)$ and the smoothness of the coefficients $p(x)$, $q(x)$, and $\rho(x)$

A detailed exposition of the methods of Ritz, Galerkin, and others can be found in the book by S.G.Mikhlin Variatsionnye metody v matematicheskoi fizike (Variational methods in mathematical physics) Gostekhizdat, 1957.

6. Basis of the Fourier method for solving the mixed problem in the case of several independent variables. The Fourier method can be used to solve the mixed problem for a hyperbolic equation of the form

$$\frac{\partial^2 u}{\partial t^2} = \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left[a_{ij}(x_1, \ldots, x_n)\frac{\partial u}{\partial x_j}\right] - a(x_1, \ldots, x_n)\,u + f(t, x_1, \ldots, x_n) \tag{26.13}$$

in a right cylinder C_T of arbitrary height T, one base of which is a region G in the hyperplane $t = 0$ under the initial conditions

$$u(0, x) = \varphi(x), \quad \frac{\partial u}{\partial t}(0, x) = \psi(x) \tag{26.14}$$

and the boundary condition

* See, for example KRYLOV, N.M. and BOGOLYUBOV, N.N., *Izv. Akad. Nauk SSSR, seriya fiz.-matem.*, 43-71 and 105-114 (1930).

$$u = 0 \text{ on the boundary of } G. \tag{26.15}$$

As in the case of two independent variables, the solution of this problem is formally represented by a series

$$\sum_{k=1}^{\infty} \left[A_k \cos\sqrt{\lambda_k}t + \frac{B_k}{\sqrt{\lambda_k}} \sin\sqrt{\lambda_k}t + \frac{1}{\sqrt{\lambda_k}} \int_0^t f_k(\tau) \sin\sqrt{\lambda_k}(t-\tau)\,d\tau \right] v_k(x_1, \ldots, x_n), \tag{26.16}$$

where the $v_k(x_1, \ldots, x_n)$ are the eigenfunctions of the equation

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial v}{\partial x_j}\right) - av + \lambda v = 0$$

with boundary condition (26.15) and

$$f_k(t) = \int \underset{G}{\cdots} \int f(t, x_1, \ldots, x_n)\, v_k(x_1, \ldots, x_n)\, dx_1 \ldots dx_n.$$

S.L. Sobolev first considered generalised solutions of the mixed problem. He obtained so-called ergodic inequalities for solutions of equations (26.13) in the cylinder C_T. These inequalities make it possible for us to prove the convergence in mean of the series (26.16) and also of the series that are obtained by differentiating this series termwise with respect to x_i and t. They also enable us to show that the sum of the series (26.16) is a generalised solution of the mixed problem (26.13)-(26.15).

Sobolev's studies on hyperbolic equations, in which theorems proven by him on the embedding of functional spaces were systematically applied and the concepts of generalised solutions and generalised derivatives were used, had a great effect on subsequent investigations of the mixed problem.

For the nonhomogeneous wave equation with several independent variables, Kh.L. Smolitskii, by using estimates derived by him for eigenfunctions and their derivatives, proved the existence of an ordinary solution to the mixed problem.

O.A. Ladyzhenskaya has shown* that under certain condi-

* LADYZHENSKAYA, O.A., *The mixed problem for hyperbolic equations*, Gostekhizdat, (1953).

tions on the coefficients of equation (26.13), the initial function, and the boundary of the region G, the series (26.16) and the series obtained by twice differentiating it termwise with respect to x_i and t converge uniformly throughout $\bar{C}_T$. V.A. Il'in has presented another justification of the Fourier method of solving of the mixed problem (26.13)-(26.15). In this case, the assumptions regarding the boundary of the region G were reduced to a minimum*. M.A. Krasnosel'skii has proposed a general procedure for justifying the Fourier method for a broad class of problems. His procedure was based on the use of the theory of fractional powers of operators in functional spaces**.

7. Solution of the mixed problem for the general linear hyperbolic second-order equation. The mixed problem for hyperbolic equations of the form

$$\frac{\partial^2 u}{\partial t^2} = \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} a_{0i} \frac{\partial^2 u}{\partial t \partial x_i} + \sum_{i=1}^{n} b_i \frac{\partial u}{\partial x_i} + b_0 \frac{\partial u}{\partial t} + cu + f, \quad (26.17)$$

where the a_{ij}, a_{0i}, b_i, b_0, c and f are sufficiently smooth functions of $t, x_1, \ldots, x_n$, was first solved by Krzyzanski and Schauder† by means of an analytic approximation of the coefficients of the equation and of the initial and boundary conditions. Here, it was necessary either to impose stringent restrictions on the smoothness of the initial conditions or to assume that the altitude of the cylinder C_T was sufficiently small.

By the method of finite differences, O.A. Ladyzhenskaya†† showed that the mixed problem for equation (26.17) can be solved in a cylinder C_T of arbitrary altitude T under certain natural assumptions regarding the coefficients and initial conditions. She also studied the question of the existence, uniqueness, and differential properties of the generalised

* IL'IN, V.A., *Uspekhi matem. nauk*, **15**, 2, 97-154 (1960).

** See KRASNOSEL'SKII, M.A. and PUSTYL'NIK, Ye.I., *Dokl. Akad. Nauk SSSR*, **122**, No. 6, 978-981 (1958).

† KRZYZANSKI and SCHAUDER, *Studia Mathematica*, **6**, 162-189 (1936).

†† See footnote to subsection 6.

solution of the mixed problem.

The mixed problem for equation (26.17) in the cylinder C_T can be reduced to the Cauchy problem for an operator equation in some functional space. This space possesses, in particular, the property that all smooth functions belonging to it satisfy the boundary conditions given on the lateral surface of the cylinder C_T. This approach turned out to be possible also for equations and systems of a more general kind. This made it possible to prove the theorems on the existence and uniqueness of the generalised solution to mixed problems for such equations and systems by the methods of functional analysis*.

Extremely general results regarding the solvability of mixed problems for different classes of equations have been obtained by means of the apparatus of the theory of generalised functions**.

* See, for example, VISHIK, M.I. and LADYZHENSKAYA, O.A., *Uspekh matem. nauk,* **11**, 6, 41-97 (1956).

** LIONS, *Acta Mathematica,* **94**, No. 1-2, 13-153 (1955).

Chapter 3

Elliptic equations

27. INTRODUCTION

In this chapter, we shall study Laplace's equation

$$\frac{\partial^2 u}{\partial x_1^2}+\dots+\frac{\partial^2 u}{\partial x_n^2}=0. \tag{27.1}$$

as the simplest representative of elliptic equations.

The basic properties of the solutions of this equation do not depend on n. For simplicity in writing, we shall therefore examine the case in which $n=2$ without always pointing it out when analogous considerations can be applied for $n>2$. At the end of the chapter, we shall give a summary of the results known for elliptic equations of a more general type.

Steady states are described by elliptic equations. In Section 1, we saw, for example, that a steady-state temperature u in a homogeneous disc or a homogeneous solid satisfies Laplace's equation. We also saw that this equation describes the shape of a membrane that is stretched across some curve in space in its equilibrium position. The potentials of a gravitational field or of a stationary electric field also satisfy Laplace's equation at points where there are no masses or electric charges.

One of the most basic properties of the solutions of elliptic equations is their smoothness. This corresponds exactly to the fact that elliptic equations describe steady states. It is

physically clear that all original unevennesses must be smoothed out by the time a steady state is attained. In this chapter, it will be shown that all continuous solutions of Laplace's equation are analytic with respect to all independent variables. However, it would not be correct to assert that all solutions of Laplace's equation are analytic. For example, the function u defined by

$$u(x, y) = \text{Reel}\{e^{-\frac{1}{z^4}}\} \text{ when } z \neq 0, \text{where } z = x + iy,$$
$$u(0, 0) = 0.$$

satisfies the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{27.2}$$

at all values of x and y. However, it is easy to see that this function not only is not analytic in a neighbourhood of the coordinate origin but is discontinuous at the origin. Thus, to assert that u is analytic, we need to assume some initial degree of smoothness for the function u. Continuous solutions of Laplace's equation (that is, continuous functions for which the derivatives $\frac{\partial^2 u}{\partial x_i^2}$ exist and for which

$$\sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} = 0)$$

are called harmonic functions.

A typical boundary-value problem for elliptic equations is the first boundary-value problem (the Dirichlet problem), of which we spoke in Section 1: A continuous function f is given on the boundary Γ of some finite region G of the space $(x_1, \ldots, x_n)$. Find a function $u(x_1, \ldots, x_n)$ that is harmonic in G and that assumes the given values of f on Γ. The precise meaning of the words 'the function $u(x_1, x_2, \ldots, x_n)$ must assume given values on the boundary' is as follows: the function coinciding with $u(x_1, x_2, \ldots, x_n)$ within G and coinciding on the boundary with the value of the function f given there must be continuous on $\overline{G} = G + \Gamma$.

The second boundary-value problem (Neumann's problem) is the following: Consider a finite region G bounded by a surface Γ with a continuously turning tangent plane. Find a harmonic function $u(x_1, \ldots, x_n)$ that is continuous in $G + \Gamma$

and whose derivative $\frac{\partial u}{\partial n}$ in the direction of the outer normal at each point of the boundary of G is equal to the value at that point of the given function f.

In Section 1, we considered examples of physical problems that lead to the first and second boundary-value problems for Laplace's equation.

In what follows, we shall consider in detail the questions of the existence and uniqueness of a solution to these problems for Laplace's equation.

Problem. Suppose that

$$u(x_1, \quad \ldots, \quad x_n) = f(r).$$

where

$$r = \sqrt{(x_1 - x_1^0)^2 + \ldots + (x_n - x_n^0)^2},$$

and that the function $f(r)$ is defined for $r > 0$ and has a continuous second derivative. Show that, if $u(x_1, \ldots, x_n)$ is harmonic for $r > 0$, then

$$f(r) = C_1 + \frac{C_2}{r^{n-2}} \qquad \text{for} \quad n \neq 2,$$

$$f(r) = C_1 + C_2 \ln \frac{1}{r} \quad \text{for} \quad n = 2.$$

28. A PROPERTY OF EXTREME VALUES AND ITS CONSEQUENCES

1. We shall confine ourselves to a consideration of harmonic functions $u(x, y)$ of two independent variables. All the assertions proven in this section are valid for harmonic functions of an arbitrary number of independent variables and they can be proven in an analogous manner.

Lemma 1. Suppose that a continuous function $u(x, y)$ is defined inside and on the boundary of a circle of radius R and that $u(x, y)$ is harmonic at all interior points of that circle. Suppose that $u(x, y) > u(x_0, y_0)$ at all interior points, where (x_0, y_0) is some point lying on the boundary of the circle. If the function $u(x, y)$ has a derivative in direction ν at the point (x_0, y_0), where ν forms an acute angle with the inwardly directly normal, then

$$\frac{\partial u}{\partial \nu} > 0.$$

Proof: Since a harmonic function remains harmonic under a translation of the coordinate axes, we may assume that the coordinate origin is at the centre of the circle. Consider the function

$$v(x, y) = \ln\frac{1}{r} + \frac{r^2}{4R^2} - \ln\frac{1}{R} - \frac{1}{4},$$

where $r = \sqrt{x^2 + y^2}$. Since on the boundary of the circle $v = 0$ and

$$\frac{\partial v}{\partial r} = -\frac{1}{r} + \frac{r}{2R^2} < 0,$$

if $0 < r \leqslant R$, we have $v > 0$ at all internal points of the circle except the centre. It is easy to show that

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = \frac{1}{R^2},$$

if $x^2 + y^2 \neq 0$.

Let us denote by D the set of points (x, y) for which $\frac{1}{4}R^2 < x^2 + y^2 < R^2$. Let α denote the smallest value of the function $u(x, y) - u(x_0, y_0)$ on the circle $x^2 + y^2 = \frac{R^2}{4}$. It follows from the conditions of the lemma that $\alpha > 0$.

Consider the function

$$w(x, y) = u(x, y) - u(x_0, y_0) - \frac{\alpha}{v\left(\frac{R}{2}, 0\right)} v(x, y)$$

defined on D. It is easy to see that $w \geqslant 0$ on the boundary of D. The function $w(x, y)$ cannot have its lowest value within the region D because

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = -\frac{\alpha}{v\left(\frac{R}{2}, 0\right) R^2} < 0$$

in D and at a minimum it is necessary that $\frac{\partial^2 w}{\partial x^2} \geqslant 0$ and $\frac{\partial^2 w}{\partial y^2} \geqslant 0$. Therefore, w must be nonnegative at all points in the region D; that is,

$$u(x, y) - u(x_0, y_0) \geqslant \frac{a}{v\left(\frac{R}{2}, 0\right)} v(x, y).$$

At the point (x_0, y_0), we have

$$\frac{\partial v}{\partial \nu} = \frac{\partial v}{\partial r} \cos(\nu, r),$$

where $\cos(\nu, r)$ denotes the cosine of the angle between the direction of the radius vector at the point (x_0, y_0) and the direction ν. Obviously $\frac{\partial v}{\partial y} > 0$. Since the functions $u(x, y) - u(x_0, y_0)$ and $v(x, y)$ vanish at the point (x_0, y_0) and since

$$u(x, y) - u(x_0, y_0) \geqslant \frac{a}{v\left(\frac{R}{2}, 0\right)} v(x, y),$$

in the region D, we have

$$\frac{\partial u}{\partial \nu} \geqslant \frac{a}{v\left(\frac{R}{2}, 0\right)} \frac{\partial v}{\partial \nu} > 0,$$

at the point (x_0, y_0), which completes the proof.

Extreme-value theorem. A nonconstant harmonic function $u(x, y)$ cannot attain either the least upper bound or the greatest lower bound of its values in G at any interior point of that region.

(If the region G is finite and if $u(x, y)$ can be extended to G in such a way that this extension, which we still denote by $u(x, y)$ will be continuous in $\overline{G}$, then it is obvious that the least upper and greatest lower bounds of the values of $u(x, y)$ in G will coincide respectively with its maximum and minimum values in $\bar{G}$.)

Proof*: Let us suppose that a nonconstant harmonic function assumes in the region G a value m equal to the greatest lower bound of the values of $u(x, y)$ in G. Suppose that E is the set of those points in G at which $u(x, y) = m$. Since the function $u(x, y)$ is nonconstant in G, there exists a region G_1 the closure of which is contained in G and which holds certain points of the set E and at least one point not belonging to E. Within the region G_1 there exists a point P not belonging to E whose distance

* OLEINIK, O.A., *Matem. sbornik*, **30** (72), 3, 696-697 (1952).

from the set E is less than the distance to the boundary of G since there exists points in G_1 that are arbitrarily close to E but that do not belong to E, and for such points in G_1, the distance to the boundary of G is greater than some positive number*.

Let us consider a circle K with centre at the point P and with radius equal to the distance from the point P to be set E. This circle lies within G, and all of its interior points belong to E. On the boundary of K, there is a point Q belonging to E. This follows from the definition of the distance from a point to a set and from the fact that the cluster points of the set E that lie in G belong to E. The derivatives $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ must vanish at points of the set E.

By applying Lemma 1 to the function $u(x, y)$ and the circle K, we see that the derivative of $u(x, y)$ at the point Q in an arbitrary direction not tangential to the circle at the point Q, if such a derivative exists, is nonzero. This contradicts the fact that $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ vanish at the point Q since at least one of the coordinate axes does not coincide with the tangent to the circle at the point Q. This contradiction shows that the harmonic function $u(x, y)$, which is not constant, cannot assume a value equal to m in G

If $u(x, y)$ assumed within G a value M equal to the least upper bound of the values of $u(x, y)$ in G, the function $-u(x, y)$ would take a value equal to the greatest lower bound of its values in G, which is impossible. This proves the theorem.

Corollary. A function that is harmonic in a finite region G and continuous in $\bar{G}$ takes its greatest and lowest values on the boundary of that region.

2. Uniqueness of the solution of the Dirichlet problem follows immediately from the theorem that we have just proven. Let us suppose that two harmonic functions u_1 and u_2 coincide on the boundary on some finite region G. Then, their difference, which obviously is also a harmonic function, is identically equal to zero on the boundary of the region, and, from what we have just shown, it cannot assume values either

* The term 'distance from the point P to the set $\mathfrak{M}$ refers to the greatest lowest bound of the distances from P to points in $\mathfrak{M}$.

greater or less than zero within the region. Thus,

$$u_1 - u_2 \equiv 0 \text{ and } u_1 \equiv u_2.$$

It also follows from the extreme-value theorem that the solution to the Dirichlet problem depends continuously on the boundary conditions for an arbitrary bounded region G. To see this, let us suppose that u_1 and u_2 are solutions to the Dirichlet problem for some region G, that the values of these solutions on the boundary Γ of the region G are given by functions f_1 and f_2, and that $|f_1 - f_2| < \varepsilon$ everywhere on Γ. Then, the boundary values of the harmonic function $u_1 - u_2$, that is $f_1 - f_2$, satisfy the inequalities

$$-\varepsilon < f_1 - f_2 < \varepsilon.$$

It then follows from the extreme-value theorem that

$$-\varepsilon < u_1 - u_2 < \varepsilon,$$

throughout the region G; that is $|u_1 - u_2| < \varepsilon$, which was to be proven.

From this, we obtain the following lemma, which will be useful in what follows.

Lemma 2. Suppose that a sequence of functions that are continuous in some closed bounded region and harmonic within that region converges uniformly on the boundary of that region. Then, it also converges uniformly throughout the entire region.

To prove this, let us consider such a sequence $u_1, \ldots, u_n, \ldots$ and let us denote by f_i the values of the function u_i on the boundary Γ of the region G. By hypothesis, the sequence $\{f_i\}$ converges uniformly. From the Cauchy criterion, for every $\varepsilon > 0$ there exist an N such that for $n, m > N$ we have $|f_n - f_m| < \varepsilon$ everywhere on Γ. But then, from what we have proven, $|u_n - u_m| < \varepsilon$ everywhere in $\bar{G}$ for these values of n and m. From the sufficiency of the Cauchy criterion, we conclude that the sequence $u_1, \ldots, u_n, \ldots$ converges uniformly in the closed region.

3. By using Lemma 1 and the extreme-value theorem, we may now prove the following theorems:

Theorem 1. Suppose that the boundary Γ of a region G is such that each point P of the boundary Γ can be touched by a circle K_P contained in the region G. In other words, suppose that, for every point P, there exists a circle K_P that holds the point P all the interior points of which belong to G (as will be the case, for example, where the curve bounding the region G has at every point a finite curvature). If a harmonic function $u(x, y)$ is continuous in $G \cup \Gamma$ and is not constant, then at the point P_1 of the boundary of G at which $u(x, y)$ assumes its smallest (greatest) value, the derivative $\frac{\partial u}{\partial n}$ of the function $u(x, y)$ in the direction of the outer normal is negative (positive) provided the derivative $\frac{\partial u}{\partial n}$ exists at that point.

Proof: Consider a circle K_{P_1}. By hypothesis, all interior points of this circle belong to G. If $u(x, y)$ is not constant, the extreme-value theorem tells us that the function $u(x, y)$ assumes its smallest value only at points on the boundary of G. Therefore, the value of $u(x, y)$ at all interior points K_{P_1} is strictly less than the value of $u(x, y)$ at the point P_1. By applying Lemma 1 of this section, we obtain $\frac{\partial u}{\partial n} < 0$ at the point P_1 provided this derivative exists.

At points of Γ at which $u(x, y)$ assumes its greatest value, the function $-u(x, y)$ assumes its smallest value and, from what has been proven, $\frac{\partial}{\partial n}(-u) < 0$ and hence $\frac{\partial u}{\partial n} > 0$.

Theorem 2. Two solutions of the same second boundary-value problem can differ from each other only by a constant term if the boundary of G satisfies the condition stated in Theorem 1.

Proof: Suppose that $u_1(x, y)$ and $u_2(x, y)$ are harmonic functions in G, that they are continuous on $G + \Gamma$, and that $\frac{\partial u_1}{\partial n} = \frac{\partial u_2}{\partial n} = f$ on Γ. The function $u(x, y) = u_1(x, y) - u_2(x, y)$, which is harmonic in G, is continuous in $G + \Gamma$ and $\frac{\partial u}{\partial n} = 0$. If $u(x, y)$ were not constant, the derivative $\frac{\partial u}{\partial n}$ would, according to Theorem 1, be nonzero at the point P_1 at which

$u(x, y)$ attains its smallest value. Consequently, $u(x, y)$ is constant.

Problem. The third boundary-value problem consists in finding a function $u(x, y)$ that is harmonic in a region G and continuous in $G \cup \Gamma$ such that $\frac{\partial u}{\partial n} + au$ assumes at every point of the boundary of the region G the value of a given function f, where $a \geqslant 0$, $a \not\equiv 0$, and $\frac{\partial u}{\partial n}$ is the derivative in the direction of the outer normal. Prove the uniqueness of the solution of the third boundary-value problem for Laplace's equation under the assumption that the boundary Γ of the region G satisfies the condition stated in Theorem 1.

29. SOLUTION OF THE DIRICHLET PROBLEM FOR A CIRCLE

1. Suppose that a continuous function $f(s)$ is given on a circle of radius 1. Here, s denotes the arc length as measured from some fixed point, and we assume that $f(0) = f(2\pi)$. Let us construct a function u that is harmonic in the interior of the circle and that takes the given values of $f(s)$ on the circle itself.

We place the coordinate origin in the centre of the circle and we let the positive path of the x-axis pass through the point $s = 0$. Let us shift to polar coordinates, taking the x-axis as the polar axis and the origin O as the pole. Then, the equation of the bounding circle in polar coordinates is $\rho = 1$, and $f(s) = f(\varphi)$, where φ is the polar angle of the point in question on the circle.

Let us apply the Fourier method to the solution of this problem, assuming first that $f(s)$ has a continuous second-order derivative. Later, we shall dispense with this restriction. Laplace's equation in polar coordinates takes the form

$$\frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial u}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 u}{\partial \varphi^2} = 0. \tag{29.1}$$

Let us seek a solution to this problem in the form of a product of two functions:

$$u = R(\rho) \cdot \Phi(\varphi). \tag{29.2}$$

When we substitute this product into equation (29.1) and separate the variables, we obtain (by procedure analogous

to that used in Section 20) for the functions $R(\rho)$ and $\Phi(\varphi)$ the ordinary differential equations

$$\Phi'' + \lambda\Phi = 0$$

$$\rho^2 R'' + \rho R' - \lambda R = 0. \tag{29.3}$$

From the nature of the problem, the function $\Phi(\varphi)$ must be periodic with period 2π, which can be the case only for λ equal to the square of an integer n. Setting $\lambda = n^2$ and

$$\Phi_n = a_n \cos n\varphi + b_n \sin n\varphi,$$

we obtain from (29.3) the following equation for R:

$$\rho^2 R'' + \rho R' - n^2 R = 0.$$

This equation* has the linearly independent solutions $R = \rho^n$ and $R = \rho^{-n}$. Since the second solution has a discontinuity at the origin, the particular solutions of the form (29.2) that are continuous inside a unit circle will be the functions

$$u_n(\rho, \varphi) = \rho^n (a_n \cos n\varphi + b_n \sin n\varphi).$$

Furthermore, for $\lambda = 0$, we obtain the solution $u_0(\rho, \varphi) = \text{const}$, which we denote by $\frac{a_0}{2}$. For arbitrary bounded a_n and b_n, the series

$$u(\rho, \varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \rho^n (a_n \cos n\varphi + b_n \sin n\varphi) \tag{29.4}$$

converges at any interior point of the circle since, for $\rho < 1$ this series is majorised by a convergent series of the form

$$M(1 + \rho_1 + \rho_1^2 + \dots + \rho_1^n + \dots), \tag{29.5}$$

where $\rho < \rho_1 < 1$.

To show that the function (29.4) is harmonic for $0 \leqslant \rho < 1$, we write the series (29.4) in Cartesian coordinates:

$$u(x, y) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \text{Reel}\,[(a_n - ib_n)(x + iy)^n]. \tag{29.6}$$

* By making the substitution $\rho = e^l$, we can reduce this equation to an equation with constant coefficients.

The series (29.6) and the series that are obtained from it by termwise differentiating it with respect to x and y an arbitrary number of times converge uniformly for $0 \leqslant \rho < \rho_1 < 1$ since these series are majorised by the series (29.5), and by the series that are obtained from (29.5) by termwise differentiating with respect to ρ_1. From this, it follows that $u(x, y)$ satisfies Laplace's equation since each term of the series (29.6) is a harmonic function.

If we set $\rho = 1$ and $u(1, \varphi) = f(\varphi)$, we obtain from (29.4) the equation

$$f(\varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\varphi + b_n \sin n\varphi),$$

which will be valid under the hypotheses made regarding $f(\varphi)$if we set a_0, a_n and b_n equal to the Fourier coefficients for the function $f(\varphi)$:

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(\varphi)\, d\varphi;$$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(\varphi) \cos n\varphi\, d\varphi;$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(\varphi) \sin n\,\varphi\, d\varphi. \tag{29.7}$$

For the series (29.4) with coefficients determined by formulae (29.7) to give the solution to the Dirichlet problem, we still need to show that this function is continuous in the closed circle $\rho \leqslant 1$ (see Section 27). Now, for $\rho \leqslant 1$ the series (29.4) is majorised by the series

$$\frac{|a_0|}{2} + \sum_{n=1}^{\infty} |a_n| + |b_n|,$$

which converges by virtue of the hypothesis that the second derivative of the function $f(s)$ is continuous*.

Thus, for twice continuously differentiable boundary functions $f(s)$, the Dirichlet problem for a unit circle is solved by the series (29.4).

* In this case, as we know, $a_n = O(1/n^2)$ and $b_n = O(1/n^2)$.

Let us now show that the series (29.4), where a_n and b_n are determined by formulae (29.7) also represents (for $\rho < 1$) the solution to the Dirichlet problem in the case in which $f(s)$ is an arbitrary continuous function. First, we construct a sequence of twice continuously differentiable functions $f_m(s)$ that converges uniformly to the continuous function $f(s)$ given on the circle $\rho = 1$. Suppose that u_m is the solution to the Dirichlet problem corresponding to the function $f_m(s)$. From Lemma 2 of Section 28, the sequence u_m converges uniformly on the circle $\rho \leqslant 1$ to the continuous function $u(x, y)$. Obviously, for $\rho = 1$, the function $u(x, y)$ coincides with $f(s)$.

Let us show that, for $\rho < 1$, the function $u(x, y)$ is given by the series (29.4) with the coefficients (29.7). From what has been shown, this series converges for $\rho < 1$ and it is a harmonic function. Suppose that $a_n^{(m)}$ and $b_n^{(m)}$ are the Fourier coefficients of the function f_m.

For sufficiently large m and all n, we have

$$|a_n - a_n^{(m)}| \leqslant \varepsilon, \quad |b_n - b_n^{(m)}| \leqslant \varepsilon,$$

where ε is an arbitrary positive number.

Therefore,

$$\left|\frac{a_0}{2} + \sum_{n=1}^{\infty} \rho^n (a_n \cos n\varphi + b_n \sin n\varphi) - u_m\right|$$
$$= \left|\frac{a_0 - a_0^{(m)}}{2} + \sum_{n=1}^{\infty} \rho^n [(a_n - a_n^{(m)}) \cos n\varphi + (b_n - b_n^{(m)}) \sin n\varphi]\right|$$
$$\leqslant 2\varepsilon \sum_{n=0}^{\infty} \rho^n = 2\varepsilon \frac{1}{1-\rho}.$$

Thus, the function $u(x, y)$ is given by the series (29.4) and is the solution to the Dirichlet problem corresponding to the harmonic function $f(s)$.

2. Let us transform the series (29.4) by replacing the coefficients a_n and b_n with the expressions given for them in formulae (29.7). For $\rho < 1$, we obtain

$$u(\rho, \varphi) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \rho^n (a_n \cos n\varphi + b_n \sin n\varphi)$$

(continued over page)

$$=\frac{1}{2\pi}\int_0^{2\pi} f(\psi)\,d\psi+\frac{1}{\pi}\sum_{n=1}^{\infty}\rho^n\left\{\int_0^{2\pi} f(\psi)\cos n\psi\,d\psi\cdot\cos n\varphi\right.$$
$$\left.+\int_0^{2\pi} f(\psi)\sin n\psi\,d\psi\cdot\sin n\varphi\right\}$$
$$=\frac{1}{2\pi}\int_0^{2\pi} f(\psi)\,d\psi+\frac{1}{\pi}\sum_{n=1}^{\infty}\int_0^{2\pi} f(\psi)\,\rho^n\cos n(\psi-\varphi)\,d\psi$$
$$=\frac{1}{2\pi}\int_0^{2\pi} f(\psi)\left(1+2\sum_{n=1}^{\infty}\rho^n\cos n(\psi-\varphi)\right)d\psi.$$

Let us set $\varphi-\psi=\omega$ in the parenthesized expression. We obtain

$$1+2\sum_{n=1}^{\infty}\rho^n\cos n\omega=-1+2\sum_{n=0}^{\infty}\rho^n\cos n\omega$$
$$=-1+2\,\mathrm{Reel}\sum_{n=0}^{\infty}\rho^n e^{in\omega}=-1+2\,\mathrm{Reel}\,\frac{1}{1-\rho e^{i\omega}}$$
$$=\frac{1-\rho^2}{1+\rho^2-2\rho\cos\omega}. \tag{29.8}$$

Therefore, for $\rho<1$,

$$u(\rho,\varphi)=\frac{1}{2\pi}\int_0^{2\pi} f(\psi)\,\frac{1-\rho^2}{1+\rho^2-2\rho\cos(\psi-\varphi)}\,d\psi. \tag{29.9}$$

The integral (29.9) is called Poisson's integral.

For a circle of arbitrary radius R and an arbitrary continuous function $f(s)$ we obtain the solution to the Dirichlet problem by replacing ρ in formula (29.9) with $\frac{\rho}{R}$. Instead of ψ, we may take $s=R\psi$ as our variable of integration. Then, we obtain Poisson's integral for an arbitrary circle

$$u(\rho,\varphi)=\frac{1}{2\pi R}\int_0^{2\pi R} f(s)\,\frac{R^2-\rho^2}{R^2+\rho^2-2R\rho\cos(\psi-\varphi)}\,ds. \tag{29.10}$$

Remark: Formulae analogous to (29.10) are valid for the solution of the Dirichlet problem for an n-dimensional

sphere. For $n=3$, the corresponding formula is of the form

$$u(r,\theta,\varphi)=\frac{1}{4\pi R}\iint\limits_{\Sigma} f(\theta',\varphi')\frac{R^2-r^2}{(R^2-2rR\cos\gamma+r^2)^{3/2}}\,d\sigma,$$

where the integration is taken over the sphere Σ of radius R, and γ is the angle between the radius vectors of the point (r,θ,φ) and the variable point (R,θ',φ') on the surface Σ.

Problem 1. Show by direct verification that Poisson's integral (29.10) is a harmonic function in a circle of radius R.

2. Show that the limiting values on the circle $\rho=R$ of the function defined by Poisson's integral (29.10) coincides with $f(s)$.

3. We shall now show a different approach to the solution of the Dirichlet problem for the circle. It is based on the use of Green's function.

We first give Green's formula, which is valid for any two functions $u(x,y)$ and $v(x,y)$, that possess continuous first and second derivatives in $D\cup\Gamma$, where D is some bounded region with a piecewise-smooth boundary Γ. If we apply Ostrogradskii's formula, we obtain

$$I=\iint\limits_D\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dx\,dy=\iint\limits_D\left[\frac{\partial}{\partial x}\left(u\frac{\partial v}{\partial x}\right)\right.$$
$$\left.+\frac{\partial}{\partial y}\left(u\frac{\partial v}{\partial y}\right)\right]dx\,dy-\iint\limits_D u\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}\right)dx\,dy$$
$$=\int\limits_\Gamma u\frac{\partial v}{\partial n}\,ds-\iint\limits_D u\,\Delta v\,dx\,dy; \quad (29.11)$$

Here, $\Delta v\equiv\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}$, and $\frac{\partial v}{\partial n}$ denotes the derivatives of v in the direction of the outer normal to the curve Γ. Analogously, we obtain

$$I=\int\limits_\Gamma v\frac{\partial u}{\partial n}\,ds-\iint\limits_D v\,\Delta u\,dx\,dy. \quad (29.12)$$

From equations (29.11) and (29.12), we get Green's formula:

$$\iint_D (v\,\Delta u - u\,\Delta v)\,dx\,dy = \int_\Gamma \left(v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n}\right) ds. \tag{29.13}$$

Suppose that the function $u(x, y)$ is harmonic in the region D. Starting with equation (29.13), we shall derive a formula giving the value of u at an arbitrary point (x_0, y_0) of the region D in terms of the boundary values of u.

We set

$$v(x, y;\ x_0, y_0) = \ln\frac{1}{r},$$

where

$$r = \sqrt{(x-x_0)^2 + (y-y_0)^2},$$

and we apply Green's formula (29.13) to the functions u and v in the region D_ε bounded by the curve Γ and the circle Γ_ε with centre at the point (x_0, y_0) and with an arbitrarily small radius ε. We obtain

$$\int_\Gamma \left[\ln\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial}{\partial n}\left(\ln\frac{1}{r}\right)\right] ds + \int_{\Gamma_\epsilon} \left[\ln\frac{1}{r}\frac{\partial u}{\partial n_\epsilon} - u\frac{\partial}{\partial n_\epsilon}\left(\ln\frac{1}{r}\right)\right] ds_\epsilon = 0. \tag{29.14}$$

The direction of the normal n_ϵ to Γ_ϵ that is directed outwardly from the region D_ϵ coincides with the direction along the radius of Γ_ϵ from the point (x, y) to the point (x_0, y_0). Therefore, on Γ_ϵ

$$\frac{\partial}{\partial n_\epsilon}\left(\ln\frac{1}{r}\right) = -\frac{\partial}{\partial r}\left(\ln\frac{1}{r}\right) = \frac{1}{r} = \frac{1}{\epsilon}. \tag{29.15}$$

Because of the continuity of the first derivatives of u, we obtain

$$\max_{\Gamma_\epsilon}\left|\frac{\partial u}{\partial n_\epsilon}\right| \leqslant C,$$

where C is independent of ε. Therefore,

$$\left|\int_{\Gamma_\epsilon} \ln\frac{1}{r}\frac{\partial u}{\partial n_\epsilon} ds_\epsilon\right| = \left|\ln\frac{1}{\epsilon}\int_{\Gamma_\epsilon}\frac{\partial u}{\partial n_\epsilon} ds_\epsilon\right| \leqslant C\cdot 2\pi\varepsilon\cdot\ln\frac{1}{\epsilon} \longrightarrow 0$$

as $\varepsilon \longrightarrow 0$.

On the basis of (29.15), we have

$$\int_{\Gamma_\epsilon} u \frac{\partial}{\partial n_\epsilon}\left(\ln\frac{1}{r}\right) ds_\epsilon = \frac{1}{\epsilon}\int_{\Gamma_\epsilon} u\, ds_\epsilon = 2\pi u(x_\epsilon, y_\epsilon),$$

where (x_ϵ, y_ϵ) is some point on Γ_ϵ. Consequently, the last integral on the left-hand side of (29.14) tends to $-2\pi\cdot u(x_0, y_0)$ as $\varepsilon \to 0$.

Taking the limit as $\varepsilon \to 0$ in equation (29.14), we obtain

$$u(x_0, y_0) = \frac{1}{2\pi}\int_\Gamma \left[\ln\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial}{\partial n}\left(\ln\frac{1}{r}\right)\right] ds. \tag{29.16}$$

Let us suppose that we have succeeded in constructing a function $v_1(x, y;\ x_0, y_0)$ that is harmonic in the region D, that has bounded first and second derivatives in D, and that coincides on Γ with the function $\frac{1}{2\pi}\ln\frac{1}{r}$. Applying Green's formula (29.13) to the functions u and v_1 in the region D, we obtain

$$0 = \int_\Gamma \left(v_1\frac{\partial u}{\partial n} - u\frac{\partial v_1}{\partial n}\right) ds. \tag{29.17}$$

If we subtract equation (29.17) from equation (29.16), we arrive at the relation

$$u(x_0, y_0) = \int_\Gamma u\frac{\partial}{\partial n}\left(-\frac{1}{2\pi}\ln\frac{1}{r} + v_1\right) ds. \tag{29.18}$$

Thus, for a function $u(x, y)$ that is harmonic in the region D, we have obtained formula (29.18), which expresses the value of this function at an arbitrary point (x_0, y_0) of the region D in terms of the values on the boundary. The function

$$G(x, y;\ x_0, y_0) = \frac{1}{2\pi}\ln\frac{1}{r} - v_1$$

is called Green's function of the Dirichlet problem for the region D. From (29.18), we have

$$u(x, y) = -\int_\Gamma f(s)\frac{\partial G}{\partial n} ds, \tag{29.19}$$

where $f(s)$ is the value of $u(x, y)$ on Γ.

In deriving the relation (29.19), we earlier assumed that

there exists a function $u(x, y)$ that is harmonic in D and that assumes the values $f(s)$ on Γ. Therefore, now that we have constructed Green's function for the region D, we still need to verify directly that the right-hand side of formula (29.19) is indeed a solution of the Dirichlet problem in the region D with the given boundary function $f(s)$.

For some regions, Green's function can be constructed in explicit form. Thus, for a circle of radius R with centre at the point O, Green's function is represented as follows:

$$G(x, y; x_0, y_0) = \frac{1}{2\pi}\left(\ln\frac{1}{r} - \ln\frac{R}{\rho r_1}\right);$$

Here,

$$r = MM_0, \ \rho = OM_0, \ r_1 = MM_1;$$

the points M and M_0 have coordinates (x, y) and (x_0, y_0) respectively, and M_1 is the point on the continuation of the radius OM_0 for which $OM_1 \cdot OM_0 = R^2$. It is easy to show (we leave this for the reader) that formula (29.19) coincides in the case of a circle with formula (29.10), which we obtained earlier by a different method.

Problem 3. Construct Green's function for the Dirichlet problem for a semicircle.

30. THEOREMS ON THE BASIC PROPERTIES OF HARMONIC FUNCTIONS

Proofs of almost all these theorems will be based on the fact that if a function u is harmonic in some closed circle $\overline{K}$, it can be represented in this circle in the form of Poisson's integral, which is very convenient for investigations. Indeed, if a function u is harmonic and, consequently, continuous in the circle $\overline{K}$, it is possible to construct from its values on this circle a function u_1 in the form of Poisson's integral that is harmonic within K and that assumes on the boundary of K the same values as does the function u. But from the theorem of the existence of a solution to the Dirichlet problem, u_1 must be identically equal to u; that is, Poisson's integral represents the original function u.

Theorem 1 (on the arithmetic mean). Suppose that a function $u(x, y)$ is harmonic inside a circle K and continuous on $\overline{K}$. Then, its value at the centre of K is equal to the arithmetic mean of its values on the circle.

Proof: Let us represent u within K according to formula (29.10). When we apply this formula to the centre, that is, for $\rho=0$, we obtain

$$u(0, \varphi)=\frac{1}{2\pi R}\int_0^{2\pi R} f(s)\,ds=\frac{1}{2\pi R}\int_0^{2\pi R} u(R, \psi)\,ds, \qquad (30.1)$$

where $\psi=\frac{s}{R}$, which means that $u(0, \varphi)$ is equal to the arithmetic mean of the values of u on the circle of radius R.

Problem 1. Prove the extreme-value theorem (Section 28) by using the above theorem for harmonic functions.

2. Suppose that a function $u(x, y)$ is continuous in a region G and that its value at the centre of an arbitrary circle contained in G is equal to the arithmetic mean of its values on the circle. Show that the function $u(x, y)$ is a harmonic function.

Theorem 2. Suppose that a function $u(x, y)$ is harmonic and bounded within a circle K. Then its value at the centre of K is equal to the arithmetic mean of its values in the interior of this circle.

Proof: Suppose that $0<R<R_0$, where R_0 is the radius of the circle K. From (30.1), we have

$$2R\cdot u(0, \varphi)=\frac{1}{\pi}\int_0^{2\pi R} u(R, \psi)\,ds.$$

When we integrate this equation with respect to R from 0 to R_0, we obtain

$$R_0^2\cdot u(0, \varphi)=\frac{1}{\pi}\int_0^{R_0} dR\int_0^{2\pi R} u(R, \psi)\,ds,$$

so that

$$u(0, \varphi)=\frac{1}{\pi R_0^2}\iint_K u(R, \psi)\,d\Omega,$$

which completes the proof.

Theorem 3. Every harmonic function $u(x, y)$ is analytic

with respect to x and y.

This means that the function $u(x, y)$ can be expanded in a series of powers of $x - x_0$ and $y - y_0$ in the neighbourhood of any point (x_0, y_0) within the region in which $u(x, y)$ is harmonic.

Proof: Suppose that a function $u(x, y)$ is harmonic in a circle K of radius R with centre at (x_0, y_0). By a translation of the coordinate origin and a similarity transformation, we can have the point (x_0, y_0) at the coordinate origin and the radius of K equal to 1. Therefore, we may assume that $x_0 = y_0 = 0$ and that $R = 1$.

In subsection 1 of Section 29, it was shown that $u(x, y)$ can be represented by the series (29.6). Let us now consider the series

$$\frac{a_0}{2} + \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \text{Reel}\left[(a_{k+l} - ib_{k+l})\, C_{k+l}^{l}\, i^l x^k y^l\right], \tag{30.2}$$

where the C_{k+l}^{l} are the binomial coefficients and $k^2 + l^2 \neq 0$. Since $C_{k+l}^{l} < 2^{k+l}$ and since a_n and b_n are bounded, the series (30.2) is majorised for $|x| < \frac{1}{2}$ and $|y| < \frac{1}{2}$ by the convergent series

$$M\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} 2^{k+l}|x|^k|y|^l,$$

where M is a positive constant.

Since the partial sums of the series (29.6) form a subsequence of partial sums of the absolutely convergent series (30.2) and since the series (29.6) converges to $u(x, y)$ the series (30.2) also converges to $u(x, y)$. Thus, we have shown that $u(x, y)$ can be expanded in a series of powers of x and y in a neighbourhood of the point $x = y = 0$.

Theorem 4 (on uniformly convergent sequences of harmonic functions). Suppose that a sequence of functions $u_k(x, y)$ (for $k = 1, 2, \ldots$) that are harmonic inside a finite region G and continuous in $\overline{G}$ converges uniformly on the boundary of G. Then, it converges uniformly throughout the entire region G. Also, the limit function will be harmonic inside G.

Proof: According to Lemma 2 of Section 28, the sequence of functions $u_n(x, y)$ converges uniformly throughout the entire region G. It remains to show that the limit function is harmonic within G. To do this, let us take a point Q within G and let us construct a circle K with centre at the point Q and contained in G. In this circle, each of the functions $u_n(x, y)$ can be represented in the form of Poisson's integral.

Suppose that

$$u_n(x, y) = \frac{1}{2\pi}\int_0^{2\pi} f_n(\psi)\,\frac{R^2-\rho^2}{R^2+\rho^2-2\rho R\cos(\varphi-\psi)}\,d\psi, \quad (30.3)$$

where the $f_n(\psi)$ are the values of u_n on the boundary of the circle K of radius R. Because of the uniform convergence of the sequence of the functions $f_n(\psi)$ and the convergence of u_n at an arbitrary interior point (x, y) of the circle K, we may pass to the limit on both sides of equation (30.3). Denoting the limit functions respectively by $u(x, y)$ and $f(\psi)$, we obtain

$$u(x, y) = \frac{1}{2\pi}\int_0^{2\pi} f(\psi)\frac{R^2-\rho^2}{R^2+\rho^2-2R\rho\cos(\varphi-\psi)}\,d\psi.$$

From this it is clear that $u(x, y)$ is harmonic in the circle K.

This theorem is often referred to as the first Harnack theorem.

Remark: It follows from this theorem that the set of generalised functions for Laplace's equation as defined at the beginning of Section 9 coincides with the class of all harmonic functions if we consider only the continuous solutions of Laplace's equation.

Theorem 5 (on monotonic sequences of harmonic functions). Suppose that a sequence of functions $u_n(x, y)$ that are harmonic in a region G converges at some interior point A of this region and that, for arbitrary n,

$$u_{n+1}(x, y) \geqslant u_n(x, y)$$

at all points of the region G. Then, the sequence $u_n(x,y)$ converges throughout the region G to some harmonic function $u(x, y)$. Furthermore, this convergence will be uniform in every closed bounded subregion of the region G.

Proof: Let us first show that our sequence converges uniformly at every point in a circle K_1 of radius R with centre at A if its closure $\overline{K}_1$ is contained in G. Let us estimate the difference $u_{n+p} - u_n = v_{n,p}$, where p is an arbitrary positive integer. By hypothesis, $v_{n,p} \geqslant 0$. Let us take a circle K^* that is concentric with K_1 but larger, having radius $R+\varepsilon$ but still contained (together with its boundary) in G. Let us represent each of the functions $v_{n,p}$ in the circle K_1 in the form of Poisson's integral

$$v_{n,p}(\rho, \varphi) = \frac{1}{2\pi}\int_0^{2\pi} v_{n,p}(R+\varepsilon, \psi)\frac{(R+\varepsilon)^2 - \rho^2}{(R+\varepsilon)^2 + \rho^2 - 2(R+\varepsilon)\rho\cos(\varphi - \psi)}\, d\psi. \tag{30.4}$$

Since $-1 \leqslant \cos(\varphi - \psi) \leqslant +1$, we have

$$\frac{R+\varepsilon-\rho}{R+\varepsilon+\rho} \leqslant \frac{(R+\varepsilon)^2 - \rho^2}{(R+\varepsilon)^2 + \rho^2 - 2(R+\varepsilon)\rho\cos(\varphi - \psi)} \leqslant \frac{R+\varepsilon+\rho}{R+\varepsilon-\rho}. \tag{30.5}$$

Since $v_{n,p}(R+\varepsilon, \psi) \geqslant 0$, we obtain on the basis of (30.4) and (30.5)

$$\frac{1}{2\pi}\frac{R+\varepsilon-\rho}{R+\varepsilon+\rho}\int_0^{2\pi} v_{n,p}(R+\varepsilon, \psi)\, d\psi \leqslant v_{n,p}(\rho, \varphi)$$

$$\leqslant \frac{1}{2\pi}\frac{R+\varepsilon+\rho}{R+\varepsilon-\rho}\int_0^{2\pi} v_{n,p}(R+\varepsilon, \psi)\, d\psi.$$

But, from the arithmetic mean theorem,

$$\frac{1}{2\pi}\int_0^{2\pi} v_{n,p}(R+\varepsilon, \psi)\, d\psi = v_{n,p}(0, \varphi) = v_{n,p}(A).$$

Therefore,

$$\frac{R+\varepsilon-\rho}{R+\varepsilon+\rho} v_{n,p}(A) \leqslant v_{n,p}(\rho, \varphi) \leqslant \frac{R+\varepsilon+\rho}{R+\varepsilon-\rho} v_{n,p}(A). \tag{30.6}$$

From this it is clear that the sequence u_n converges uniformly in $\overline{K}_1$ if it converges at the point A. Therefore, in accordance with the first Harnack theorem, the limit function is harmonic within K_1.

To show that the sequence of the u_n converges at any point B of the region G, we connect this point with A by a

broken line l consisting of a finite number of links and lying entirely within G. This is possible because of the definition of the region. The broken line l including the end points A and B is a closed set. Since it has no points in common with the boundary of G, it lies at a positive distance δ from this boundary, which is also a closed set. Now, let us take a point A_2 on the intersection of the circle K_1 and the line l. Let us describe a circle K_2 of radius $\delta/2$ with centre at A_2 From what was said above, the sequence of the u_n converges uniformly inside and on this circle. Similarly, it converges uniformly inside and on the circle K_3 of radius $\delta/2$, where the centre of K_3 lies on the intersection of l and the circle K_2. We can cover the entire line with a finite number of such circles K_i (for $i = 1, \ldots, N$) in such a way that the point B will lie within K_N. This shows that everywhere on the line l and, particularly, at the point B the sequence of the u_n converges. Since this sequence converges uniformly in each of the circles K_i and, in particular, in K_N, the limit function must, from the first Harnack theorem, be harmonic in a neighbourhood of B.

Let us now show that the sequence of the $u_n(x, y)$ converges uniformly on every closed bounded set F contained in G. From the Heine-Borel theorem, the set F can be covered by a finite number of the circles $K_1, \ldots, K_N$, whose closures are contained in G. From what we have just proven, the sequence of the $u_n(x, y)$ converges at the centre of each of these circles. Consequently, from what was proven above, this sequence converges uniformly in each of the circles K_i, and, consequently, on the entire set F.

This theorem is often called the second Harnack theorem.

Theorem 6 (bounds of the derivatives of harmonic functions). Suppose that a family of uniformly bounded harmonic functions is given in a region G. Then, in an arbitrary region G' whose closure is contained in G, the derivatives of all functions belonging to the family are uniformly bounded.

Proof: Suppose that M is the least upper bound of the absolute values of the functions belonging to the family in question and that $l > 0$ is the shortest distance from the boundary of G' to the boundary of G. Then, the circle K of radius $\frac{l}{2}$ with centre at an arbitrary point (x_0, y_0) of

the region G' lies entirely within G.

Since the derivative of a harmonic function is also harmonic, we obtain from Theorem 2

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{1}{\pi\left(\frac{l}{2}\right)^2} \iint\limits_K \frac{\partial u}{\partial x} dx dy = \frac{4}{\pi l^2} \int\limits_S u \cos(n, x) ds. \tag{30.7}$$

Here, $u(x, y)$ is an arbitrary function belonging to the family in question, S is the boundary of the circle K, and n is the outer normal to S. From (30.7), we have the inequality

$$\left|\frac{\partial u}{\partial x}(x_0, y_0)\right| \leqslant \frac{4}{\pi l^2} M \cdot 2\pi \frac{l}{2} = \frac{4M}{l}.$$

Since the point (x_0, y_0) and the function u are arbitrary, it follows that the derivatives with respect to x of functions belonging to the family are uniformly bounded in G'. The uniform convergence of their derivatives with respect to y in G' is proven analogously.

Theorem 7 (on the compactness of the family of uniformly bounded harmonic functions). An arbitrary infinite collection of harmonic functions that are uniformly bounded in a region G contains an infinite sequence that converges uniformly in an arbitrary bounded region G' whose closure is contained in G.

This assertion follows from a theorem of Arzelà since the set consisting of all these is, as a consequence of Theorem 6, equicontinuous in G'.

Theorem 8 (Liouville). A function $u\ (x, y)$ that is harmonic on the entire xy-plane cannot have either an upper or a lower bound unless it is a constant.

Proof: Suppose, for example, that $u(x, y)$ is always greater than some constant M. By adding a constant to the function $u(x, y)$ if necessary, we can always have $M \geqslant 0$. Let us assume that $M \geqslant 0$, and let us show that the value of u at an arbitrary point $Q(\rho, \varphi)$ is exactly equal to the value of u at the coordinate origin (pole) O. This will show that u is constant. To do this, we take a circle K with centre at the point O and with radius R sufficiently large that the point $Q(\rho, \varphi)$ lies inside it. When we re-

present the function u in K in the form of Poisson's integral, we obtain

$$u(Q) = \frac{1}{2\pi}\int_0^{2\pi} u(R, \psi)\frac{R^2 - \rho^2}{R^2 + \rho^2 - 2R\rho\cos(\varphi - \psi)}\,d\psi.$$

From this, we obtain in analogy with (30.6)

$$\frac{R-\rho}{R+\rho}u(O) \leqslant u(Q) \leqslant \frac{R+\rho}{R-\rho}u(O).$$

As R increases without bound, we have $u(O) \leqslant u(Q) \leqslant u(O)$, so that $u(Q) = u(O)$.

Since Q is an arbitrary point in the plane, this means that the function u is constant.

Theorem 9 (on isolated singularities). Suppose that $u(x, y)$ is a bounded function that is harmonic in a neighbourhood of a point A except at the point A itself (where $u(x, y)$ is undefined). Then, the function $u(x, y)$ can be defined at the point A in such a way that $u(x, y)$ will be harmonic throughout the entire neighbourhood of A including the point A itself.

Proof: For simplicity of notation, we take as the point A the coordinate origin. Suppose that K is a circle of radius R with centre at A and that K lies within the neighbourhood in question of A. Suppose that u_1 is a function that is harmonic within K and that coincides with u on the boundary of K. Let us set $u - u_1 \equiv v$. The function $v(x, y)$ will be bounded and harmonic throughout the circle K, except at the point A, at which it is not defined. On the circle K, the function v vanishes. Let us show that $v \equiv 0$ (and hence that $u = u_1$ everywhere in the interior of K except at the point A. Once this is done, we shall set the function v equal to zero at the point A. This will make $u = u_1$ everywhere in the interior of K, which will complete the proof of the theorem.

To prove the identity $v \equiv 0$ throughout the entire circle K except at the point A, let us define the function

$$w_\varepsilon(P) = \frac{M\ln\frac{\rho}{R}}{\ln\frac{\varepsilon}{R}},$$

everywhere in the circle, where M is the least upper bound of $|v|$ in K, ε is some small positive number, and $\rho = AP$. The function $w_\varepsilon(P)$ is a harmonic function in the region bounded by the circles $\rho = R$ and $\rho = \varepsilon$, and it vanishes for $\rho = R$ and is equal to M for $\rho = \varepsilon$. From the extreme-value theorem of harmonic functions, we have, for an arbitrary point P lying in the annulus between the circles $\rho = R$ and $\rho = \varepsilon$ (for any ε)

$$|v(P)| \leqslant M \frac{\ln \frac{\rho}{R}}{\ln \frac{\varepsilon}{R}}, \tag{30.8}$$

since $-w_\varepsilon(P) \leqslant v(P) \leqslant w_\varepsilon(P)$ on these circles. But the right member of inequality (30.8) approaches 0 as ε approaches 0. Therefore, the left member is equal to 0 since it is independent of ε.

Remark 1: Theorem 9 remains valid under the more general formulation: Suppose that $u(x, y)$ is a harmonic function in a neighbourhood of the point A except at the point A (at which $u(x, y)$ is underfined) and that, for an arbitrary point P in this neighbourhood,

$$|u(P)| \leqslant \mu(P) \ln \frac{1}{AP}, \tag{30.9}$$

where $\mu(P) \to 0$ as $P \to A$. Under these conditions, the functions $u(x, y)$ can be defined at the point A in such a way that $u(x, y)$ will be harmonic throughout the entire neighbourhood of the point A including at the point A itself.

Proof of this proposition is analogous to the proof of Theorem 9.

2: Suppose that $u(x, y)$ is bounded and harmonic in a region G and that it is continuous at all but a finite number of points on the boundary of G. Under these conditions, the function $u(x, y)$ cannot take values within G that are greater than the least upper bound of the values of $u(x, y)$ on the boundary of G or lower than the greatest lower bound of the values of $u(x, y)$ on the boundary of G.

To show this, suppose that M is the least upper bound of the values of $u(x, y)$ on the boundary of G. For simplicity, let us assume that $u(x, y)$ is continuous at all points on the boundary of G with the exception of one point P_1. Suppose that all points of G lie at a distance from P_1 no greater than R. Let us define the function

$$w_\varepsilon(P) = M + \varepsilon \ln \frac{R}{P_1P}.$$

Consider the region G_δ consisting of the points of the region G whose distance from P_1 exceeds δ. It is easy to see that $u(P) < w_\varepsilon(P)$ on the boundary of this region if δ is sufficiently small. From the extreme-value theorem of harmonic functions, $u(P) < w_\varepsilon(P)$ in G_δ. When we let ε approach zero, we have $u(P) \leqslant M$ at any arbitrary point P in the region G. In just the same way, we see that $u(P) \geqslant m$, where m is the greatest lower bound of the values $u(x, y)$ on the boundary of the region G.

Remark 3: All the properties of harmonic functions of two independent variables that we have proven in the present section remain valid for harmonic functions of an arbitrary number of independent variables and can be proven in an analogous manner. Condition (30.9) in the case of $n > 2$ independent variables should be replaced with the condition

$$|u(P)| \leqslant \mu(P) \frac{1}{(AP)^{n-2}},$$

where AP is the distance from the point A to the point P and where $\mu(P)$ approaches zero as P approaches A.

31. PROOF OF THE EXISTENCE OF A SOLUTION TO THE DIRICHLET PROBLEM

The idea of the proof that we are about to give belongs to Poincaré. Poincaré's original proof was somewhat improved by Perron. Since the following considerations are equally applicable to regions in a space of an arbitrary number of dimensions, we shall not confine ourselves to a consideration of the two-dimensional case.

1. Basic definitions and method of solving the problem. Suppose that a continuous function v is defined in an n-dimensional bounded region G and on its boundary. We shall always denote by K any n-dimensional sphere all interior points of which belong to G. We shall denote by $(v)_K$ a continuous function that is equal to v outside K and on its boundary and which is harmonic inside this sphere K. For the function v to be harmonic, it is obviously necessary and sufficient that $(v)_K \equiv v$ for every sphere K. (Of

course, for $n=2$, the 'sphere' is a circle.)

We shall call a function v superharmonic (resp. subharmonic) if, for every sphere K,

$$(v)_K \leqslant v \,(\text{resp.}\, (v)_K \geqslant v). \tag{31.1}$$

Suppose that a continuous function f is defined on the boundary of G and that a superharmonic (resp. subharmonic) function v is defined in G with the property that, on the boundary of G,

$$v \geqslant f \quad (\text{resp. } v \leqslant f).$$

We shall call this function v a superfunction (resp. subfunction) for the function f.

In all that follows, we shall consider only superharmonic and subharmonic superfunctions and subfunctions that are continuous within G and on its boundary. Therefore, when we speak of super- and subharmonic functions, we shall assume that they are continuous within and on the boundary of G without explicitly mentioning this fact.

The Poincaré-Perron method consists in the following: For a given bounded region G and a continuous function f defined on its boundary, we define the collection of all superfunctions. It is clear that this collection is not empty since every constant $c \geqslant \sup f$ is a superfunction. Let us define the value of the function u at a point P belonging to $\overline{G}$ as the greatest lowest bound of the set of values of all superfunctions at this point. We shall show that the function u is harmonic within G, that it assumes the given values of f and is continuous at all boundary points of this region at which certain conditions of which we shall speak later are satisfied. As a preliminary, we need to prove certain properties of superharmonic functions and superfunctions.

2. *Certain properties of superharmonic functions and superfunctions.*

Theorem 1. (a) Every harmonic function is superharmonic and subharmonic. (b) If v is superharmonic and u is harmonic, then $v \pm u$ is superharmonic. (c) The sum of two (and hence of an arbitrary number of) superharmonic functions is superharmonic. (d) If v is superharmonic and w is subharmonic, then $v - w$ is superharmonic.

Analogous theorems hold for subharmonic functions.

The first of the above assertions is obvious. The other three are easily proven if we keep in mind the fact that

$$(v_1 + v_2)_K = (v_1)_K + (v_2)_K.$$

With the aid of this relationship, we shall prove assertion (c). Suppose that v_1 and v_2 are two superharmonic functions. Then,

$$(v_1)_K \leqslant v_1, \quad (v_2)_K \leqslant v_2,$$

and consequently,

$$(v_1 + v_2)_K = (v_1)_K + (v_2)_K \leqslant v_1 + v_2,$$

that is, $v_1 + v_2$ is a superharmonic function.

Theorem 2. A function v that is superharmonic in a region G assumes its smallest value on the boundary of G.

Proof: Let us suppose that the function v assumes its smallest value m at some point P belonging to the interior of G. Then, we can draw a sphere K with P as centre that touches the boundary G. On the boundary of this sphere, v must everywhere be equal to m because otherwise we would have, from the arithmetic mean theorem,

$$v < (v)_K.$$

at the point P. Consequently, the function v must be equal to m at certain points of the boundary G.

Theorem 3. Every superfunction v is everywhere at least as great as any subfunction w.

Proof: According to Theorem 2, the superharmonic function $v - w$ assumes its smallest value on the boundary of the region, but it is nonnegative there. Consequently, it is nonnegative in the interior of the region.

Theorem 4. The function

$$v = \min\{v_1, v_2, \ldots, v_n\},$$

where $v_1, v_2, \ldots, v_n$ are superfunctions, is also a superfunction.

Proof: It is clear that the function v is continuous within G and on its boundary and that $v \geqslant f$ on the boundary of G. It remains to show that v satisfies inequality (31.1) for every sphere K. We note that $v(P)$ is equal to the value at the point P of one of the functions $v_1, v_2, \ldots, v_n$, let us say, v_1. Therefore, at the point P,

$$v = v_1 \geqslant (v_1)_K \geqslant (v)_K,$$

which proves inequality (31.1). Here, we use the fact that $v \leqslant v_1$ implies that everywhere $(v)_K \leqslant (v_1)_K$.

Theorem 5. If v is a superfunction, then $(v)_K$ is also a superfunction.

Proof: Let us set

$$(v)_K = z.$$

Of all the properties that a superfunction must satisfy, the only one that needs to be proven is, obviously, that for any sphere K_1,

$$(z)_{K_1} \leqslant z. \tag{31.2}$$

This property is also obvious if the sphere K_1 lies entirely within K or entirely outside K. It remains only to consider the case in which the sphere K lies within K_1 or when the boundaries of these spheres intersect.

On the boundary of K_1,

$$z \leqslant v.$$

Therefore, even within K_1,

$$(z)_{K_1} \leqslant (v)_{K_1},$$

since both the functions $(z)_{K_1}$ and $(v)_{K_1}$ are harmonic within K_1. Since v is superharmonic, we have

$$(v)_{K_1} \leqslant v.$$

Outside the sphere K and on its boundary, the functions z and v coincide. Therefore, whether K is within K_1 or whether they intersect, relation (31.2) is valid outside K and on it. The validity of this same relation in the intersection KK_1 of the interiors of K and K_1 follows from the fact that the functions z and $(z)_{K_1}$ are harmonic within KK_1, and since

relation (31.2) is valid on the boundary KK_1, it is also valid inside that region.

3. Proof that the greatest lower bound u(P) of all super functions is harmonic. To prove that the function u is harmonic in the entire region, it will obviously be sufficient to prove that it is harmonic in an arbitrary sphere K. Let us take one of the superfunctions v_1 whose value is no greater than $u(P)+\varepsilon$ at the centre P of the sphere K. We may assume that v_1 is harmonic inside K. If v_1 were not harmonic inside K, then, instead of v_1 we could take $(v_1)_K$, which, from Theorem 5, is also a superfunction and which, like v_1, assumes a value no greater than $u(P)+\varepsilon$ at the point P.

Let us also take the superfunction v_2', which assumes a value no greater than $u(P)+\frac{\varepsilon}{2}$ at the point P. We set

$$v_2 = (\min(v_1, v_2'))_K.$$

From Theorems 4 and 5, the function v_2 is also a superfunction.

Continuing these constructions, we obtain an infinite decreasing sequence of superfunctions $v_1, v_2, \ldots, v_n, \ldots$ that are harmonic inside K. This sequence is bounded below (see Theorem 3). Consequently, from Theorem 5 of Section 30 (the second Harnack theorem), this sequence converges uniformly within K to some harmonic function v.

Let us show that $v=u$ inside K. We suppose that this is not true. Then, there exists a superfunction z that assumes a value smaller than $v(P_1)$ at some point P_1 inside the sphere K. We describe a sphere K_1 of radius ρ with centre at the point P the surface of which holds the point P_1. Then, every function

$$z_n = (\min(z, v_n))_{K_1}$$

is a superfunction. Since the sequence v_n in $\overline{K}_1$ converges uniformly to v, it follows that $z_n(P)$ also converges uniformly to $\overline{K}_1$. Therefore, for sufficiently large n, the function $z_n(P)$ differs by an arbitrarily small amount from the value at the point P of the function $(\min(z, v))_{K_1}$, which is less than $v(P)$, which in turn is equal to $u(P)$. However, this contradicts the assumption that $u(P)$ is the greatest lower bound

of the values of all superfunctions at the point P.

It is customary to call the greatest lower bound of all superfunctions the generalised solution, corresponding to the given boundary function f, of the Dirichlet problem. Obviously, if the Dirichlet problem has a solution in the region G that assumes the given values of f on the boundary, this solution will coincide with the generalised solution of the Dirichlet problem corresponding to the given function f. We shall call a point Q on the boundary of G regular if, for any continuous function f defined on the boundary G the generalised solution of the Dirichlet problem corresponding to the function f is continuous at Q and is equal to $f(Q)$. Below we shall give a number of sufficient conditions for a point Q on the region G to be regular.

4. *Behaviour of a function u(P) on boundary of G.*

Theorem. A function $u(P)$ is continuous and assumes the value $f(Q)$ at a boundary point Q if this point satisfies Condition A. A superharmonic function ω_Q (called a barrier function) exists that is continuous within G and on its boundary and that possesses the following properties: (1) $\omega_Q(Q) = 0$, (2) at all points P of the region G and its boundary, except the point Q,

$$\omega_Q(P) > 0.$$

Proof: For any positive number ε, we can always, because of the continuity of f, choose a neighbourhood U_Q of the point Q sufficiently small that, at each point P, in it belonging to the boundary G, we have

$$f(Q) - \varepsilon \leqslant f(P) \leqslant f(Q) + \varepsilon.$$

Therefore, by using the fact that everywhere in $\bar{G}$ outside U_Q the function $\omega_Q(P)$ exceeds some positive constant, it is easy to show that the function

$$\varphi(P) = f(Q) - \varepsilon - C\omega_Q(P),$$

will be a subfunction provided $C > 0$ is chosen sufficiently great and that the function

$$\psi(P) = f(Q) + \varepsilon + C\omega_Q(P)$$

will be a superfunction.

Let us show, for example, that the function $\psi(P)$ is a superfunction. It is easy to see that it is superharmonic for every nonnegative C. It remains to show that it is nowhere less than f on the boundary of G. The validity of this assertion in the neighbourhood U_Q of the point Q follows from the definition of the neighbourhood U_Q and from the fact that $C\omega_Q(P) \geqslant 0$. Outside this neighbourhood, on the other hand, $\omega_Q(P)$, exceeds some positive constant by assumption and therefore, for sufficiently large C, the quantity $C\omega_Q(P)$ can be made arbitrarily large everywhere on the boundary of G outside U_Q.

Obviously, for every positive ε, the function $u(P)$ lies between these two continuous functions $\varphi(P)$ and $\psi(P)$. Consequently,

$$f(Q) - \varepsilon = \varphi(Q) \leqslant \varliminf_{P \to Q} u(P) \leqslant \varlimsup_{P \to Q} u(P) \leqslant \psi(Q) = f(Q) + \varepsilon.$$

Since ε is arbitrarily small,

$$\lim_{P \to Q} u(P) = f(Q),$$

and the function $u(P)$ that we have constructed is continuous at the point Q.

For $n > 2$, it is easiest of all to construct a barrier for a boundary point Q for which there exists an n-dimensional sphere with centre at some point O within which there is no point of the region G and whose boundary holds only one point Q in common with the boundary of G. Then, for the function $\omega_Q(P)$, we may take the function

$$\frac{1}{OQ^{n-2}} - \frac{1}{PO^{n-2}},$$

where PO (resp. OQ) denotes the distance between the points P and O (resp. O and Q). For every $n > 2$, this function is harmonic.

In the case of $n = 2$, one can show that every boundary point Q of a region bounded by a single non-self-intersecting curve satisfies Condition A. This is true because if Q is taken as the coordinate origin, the function $-\frac{p}{p^2+q^2}$, where p and q are respectively the real and imaginary parts of $\ln \frac{x+iy}{2D}$, possesses the same properties of ω_Q,

where D denotes the diameter of the region G. But the function

$$-\frac{p}{p^2+q^2} = -\text{Reel}\,\frac{1}{\ln\frac{x+iy}{2D}}$$

may no longer possess these properties if the point Q lies on the boundary of a non-simply-connected region G. This will be the case, for example, if the region G lies between two concentric circles and the point Q lies on the smaller of them. In this case, the function $-\frac{p}{p^2+q^2}$ is no longer single-valued. Therefore, it is expedient to replace Condition A with the following more general

Condition B. For an arbitrarily small neighbourhood U_Q of a point Q (here, U_Q denotes that portion of the entire neighbourhood of the point that belongs to the region G and its boundary), there exists a single-valued superharmonic function Ω_Q (a barrier function) possessing the following properties: 1. Ω_Q is defined within U_Q and on its boundary, and it is everywhere continuous. 2. $\Omega_Q(Q)=0$. 3. $\Omega_Q>0$ at every point except Q.

From these three properties of Ω_Q a fourth property can be derived:

4. $\Omega_Q \geqslant k > 0$ at all boundary points of U_Q that belong to G; here, k is some constant.

Let us show that if the point Q satisfies Condition B, it must also satisfy Condition A. Let us construct the function $\omega_Q(P)$, setting

$$\omega_Q(P) = \min\left\{\frac{2}{k}\,\Omega_Q(P),\ 1\right\} \text{inside } U_Q,$$

$$\omega_Q(P) = 1 \text{ outside } U_Q.$$

We assert that this function possesses all properties enumerated in Condition A. Specifically, (1) $\omega_Q(P)$ is continuous throughout $\overline{G}$; (2) $\omega_Q(Q)=0$; (3) $\omega_Q>0$ at all points of $\overline{G}$ except at the point Q; (4) it remains to show that the function $\omega_Q(P)$ is superharmonic, that is, that

$$(\omega_Q)_K \leqslant \omega_Q. \tag{31.3}$$

We denote by G_1 that portion of G at which $\omega_Q=1$ and we

denote by G_0 the remainder of the region G. Then, it will be obvious that (31.3) is valid if the interior of the sphere K holds either points of G_1 alone or points of G_0 alone. Therefore, we consider the only remaining possible case, namely, that the interior of the sphere K contains points belonging to G_1 and points belonging to G_0. In this case, (31.3) will be valid for that portion of the sphere K that belongs to G_1 because there $\omega_Q = 1$ and $(\omega_Q)_K \leqslant 1$. On the other hand, (31.3) is valid for points of the intersection KG_0 of the regions K and G_0 because in each of these regions from which KG_0 is constituted, the function ω_Q is superharmonic and the function $(\omega_Q)_K$ is harmonic. Furthermore, the values of ω_Q on the boundary of each such region is not less than the values of $(\omega_Q)_K$

For $n = 1$, the boundary value problem that we are considering is trivial. Therefore, in what follows, we shall assume that $n \geqslant 2$.

In the case of $n = 2$, it is easy to show that boundary point Q of the region G satisfies Condition B if the point Q is an end point of some curve l lying outside $G + \Gamma$ that intersects all neighbourhoods of sufficiently small radius with centre at the point Q. Specifically, we translate the coordinate origin to the point Q and we assume that the neighbourhood U_Q is sufficiently small that all its points lie at a distance less than c from Q, where $c < 1$, and the arc l intersects the circle enclosing U_Q. If we now set

$$\ln(x + iy) = p + iq,$$

the function

$$\Omega_Q = -\frac{p}{p^2 + q^2}$$

will possess all the properties enumerated in Condition B.

In the case of $n > 2$ it is easy to construct the function Ω_Q for every boundary point Q that is the vertex of some circular n-dimensional cone C_Q with rectilinear generators and with all points sufficiently close to Q lying in G. Let us consider a simply-connected region G^* formed of points lying inside some n-dimensional sphere S of radius R with centre at the point Q and outside the cone C_Q. On the boundary of G^*, we define the function f^* by setting

$$f^*(P) = QP,$$

where QP denotes the distance between the points P and Q. The greatest lower bound u^* of all superfunctions constructed for the region G^* and the function f^* assumes (because of the criterion formulated on page 268 concerning a sphere) the value f^* at all points of the boundary G^* with the exception of the point Q, to which this criterion is not applicable. To see that the function u^* possesses all the properties of the function Ω_Q, we still need to show that it assumes the value 0 at the point Q. To do this, we note first of all that $\underline{u}^*(Q) \geqslant 0$, since the identically zero function is a subfunction. We use the following notations:

$$\underline{u}^*(Q) = \varliminf_{P \to Q} u^*(P), \text{ and } \overline{u}^*(Q) = \varlimsup_{P \to Q} u^*(P).$$

It remains to show that

$$\overline{u}^*(Q) = 0.$$

Let us suppose that this is not the case but that

$$\overline{u}^*(Q) = c > 0. \tag{31.4}$$

We then take the point Q as the coordinate origin and we consider the function

$$u^{**}(x_1, \ldots, x_n) = u^*(kx_1, \ldots, kx_n),$$

where $k > 1$. Obviously,

$$\overline{u}^{**}(Q) = \overline{u}^*(Q) = c > 0.$$

But, on the other hand, by using the fact that $u^*(P) < R$ within G^*, it is easy to see that

$$u^* \leqslant c_1 u^{**}, \tag{31.5}$$

everywhere on the boundary of the region G^{**} at which the function u^{**} is defined except at the point Q. Here, c_1 is some constant less than unity that depends on k. The function $u^* - c_1 u^{**}$, which is harmonic in G^{**}, is continuous at all points of the boundary G^{**} except at the point Q, and the least upper bound of the values of $u^* - c_1 u^{**}$ on the boundary of G^{**} is nonpositive. Therefore, according to Remark 2 in Section 30, $u^* - c_1 u^{**} \leqslant 0$ everywhere in the region G^{**}.

The fact that relation (31.5) is satisfied throughout the

entire region G^{**} implies that

$$\bar{u}^*(Q) \leqslant c_1 \bar{u}^{**}(Q) = c_1 c.$$

Since $c_1 < 1$, this relation contradicts (31.4) if $c > 0$.

Problem 1. Show that no function $u(x, y)$ exists that is continuous on and inside the circle $x^2 + y^2 \leqslant 1$ and harmonic everywhere in the interior of this circle except at the centre and that assumes the value 0 on the circle itself and the value 1 at the centre.

Problem 2. Show that no function $u(x, y, z)$ exists that (1) is continuous in the cylinder

$$\{x^2 + y^2 \leqslant 1, \ -1 \leqslant z \leqslant 1\}$$

(2) is harmonic everywhere inside this cylinder except on the segment $-\frac{1}{2} \leqslant z \leqslant \frac{1}{2}$, $x = y = 0$, and (3) assumes on this segment the value 1 and on the boundary of the cylinder the value 0.

32. THE EXTERNAL DIRICHLET PROBLEM

We shall call the following problem the external Dirichlet problem:

Suppose that we have a bounded region G such that points not belonging to G or to its boundary Γ constitute a region whose boundary is Γ. Suppose that a continuous function f is given on the boundary of this region. Find a function $u(P)$ that is harmonic outside $G \cup \Gamma$ and that takes the values of f on Γ.

Here, we shall say that the function u assumes the values of f on the boundary of G if the function v that coincides with u outside $G \cup \Gamma$ and with f on Γ is continuous whereever it is defined.

Example. Suppose that a definite time-independent temperature u exists at every point (x, y, z) of space outside a body and on the boundary of that body. We know that in this case, u satisfies Laplace's equation outside G. Thus, to find the steady-state temperature outside G, we need to solve the external Dirichlet problem.

If we do not impose any restrictions on the behaviour of the solution to the external Dirichlet problem at distant points of space, this problem has many solutions. To ensure uniqueness of its solution, we require in the two-dimensional

case that the solution be bounded and in the many-dimensional case that the solution approach zero as the point P approaches infinity. (The function $u(P)$ is said to approach zero as $P \longrightarrow \infty$ if $|u(P)| < \varepsilon$, where ε is an arbitrary positive number, for all points P lying outside a sphere of sufficiently large radius with centre at the coordinate origin.)

The solution to the external Dirichlet problem reduces to the solution of the Dirichlet problem for bounded regions which we studied in Section 31 and which now, by analogy with the external Dirichlet problem, we shall call the internal Dirichlet problem. Here, a role is played by those supplementary conditions at infinity that are imposed on the solution of the external Dirichlet problem (more precisely, that are imposed on the values of this solution at distant points).

Inside the region G, let us take some point O and a sphere (circle in the two-dimensional case) S of radius R with centre at the point O. Let us make a transformation of the space by means of inverse radius vectors relative to this sphere; that is, let us make a transformation under which every point P of this space is mapped into a point P^* lying on a ray OP such that $OP \times OP^* = R^2$. Under this transformation, points on the sphere S remain unchanged, but the entire portion of space outside (resp. inside) S is mapped onto that portion of space lying inside (resp. outside) S.

Thus, all those points of space lying outside G are mapped into a bounded region G^* surrounding the point O. To every point in G^* except O there corresponds under the transformation one and only one point lying outside G. No point in space corresponds to the point O under this transformation. Further consideration must be made separately for two- and three-dimensional spaces.

Let us first consider the two-dimensional (plane) case. Suppose that u is the solution of the external Dirichlet problem for the region G. We set

$$u^*(P^*) = u(P) \text{ and } f^*(P^*) = f(P).$$

The function u^* will be defined everywhere in the region G^* except at the point O and will assume the value of $f^*(P^*)$ on the boundary of G^*. By direct calculation, one can show that the function $u^*(P^*)$ will be a harmonic function of the coordinates of the point P^* (more briefly, a

harmonic function of P^*) if $u(P)$ is a harmonic function of P†.

If the function $u(P)$ is bounded, $u^*(P^*)$ is also bounded. Then, from the theorem on isolated singularities, u^* can now be defined at the point O in such a way that the final function is harmonic throughout the region G^*. From the theorem on the uniqueness of the solution of the internal Dirichlet problem, it follows that a bounded function u^* is uniquely defined in G^* by its values on the boundary. It follows from this that the solution of the external Dirichlet problem is unique in the class of bounded functions. The existence of a solution follows from the fact that all points on the boundary of G^* are regular because G is simply connected (see page 269).

The three-dimensional case. Suppose again that u is the solution of the external Dirichlet problem for the region G. Let us set

$$u^*(P^*) = \frac{R}{OP^*} u(P)$$

or, equivalently,

$$u(P) = \frac{R}{OP} u^*(P^*). \tag{32.1}$$

Analogously we set

$$f^*(P^*) = \frac{R}{OP^*} f(P)$$

or, equivalently,

$$f(P) = \frac{R}{OP} f(P^*).$$

This defines the function u^* everywhere within G^* except at the point O. It will assume the value f^* everywhere on the boundary G^*. By converting the equation to spherical coordinates, we can show by direct calculations that $u^*(P^*)$ will be a harmonic function of P^* if $u(P)$ is a harmonic function of P. If $u(P)$ approaches zero as $P \longrightarrow \infty$, then, as one can easily see, $u^*(P^*)$ will satisfy the equation

$$|u^*(P^*)| \leqslant |u(P)| \frac{R}{OP^*}, \text{where } |u(P)| \longrightarrow 0 \text{ when } OP^* \longrightarrow 0.$$

† To prove this, we need to reduce Laplace's equation to polar coordinates with pole at the point O. The simplest formulas for our transfomation are written in polar coordinates.

Then, according to Remarks 1 and 3 of Section 30, the function u^* can be defined at the point O in such a way that it is harmonic inside G^*. Because the solution to the internal Dirichlet problem is unique, the bounded function u^* is uniquely determined in G^* by its values on the boundary of G^* From this it follows that the solution to the external Dirichlet problem is unique in the class of functions that tend to zero as $P \longrightarrow \infty$.

If the region G^* is such that all the points on its boundary are regular, it follows from what we have said above that the external Dirichlet problem has a solution for the region G for every continuous function defined on its boundary. Furthermore, this solution will, as can easily be seen from (32.1), satisfy the equation

$$|u(P)| \leqslant \frac{M}{OP},$$

where M is some constant and OP is the distance from the point P to some fixed point O.

Examples. When the function defined on the boundary is everywhere equal to some constant C, the function that is equal to C everywhere throughout G is the unique bounded solution to the Dirichlet problem for the plane.

In three-dimensional space, when the region in question is bounded by the sphere of radius R with centre at the point O and when the function defined on the sphere is equal to a constant C, the function

$$u(P) = \frac{C \cdot R}{OP} \tag{32.2}$$

will be the solution to the Dirichlet problem. This is the only solution of the external Dirichlet problem in question in the class of functions that approach zero as $OP \longrightarrow \infty$.

One can show that the solutions of the following two heat-flow problems tend to a constant C in the two-dimensional case and to the function (32.2) in the three-dimensional case.

1. Suppose that the temperature on the surface of an infinitely long cylindrical tube has a constant value C.

The initial temperature of the surrounding air is zero. Then, the temperature $u(t, x, y, z)$ of the air at an instant t at a point (x, y, z) approaches C as $t \longrightarrow \infty$. Physically, this means that one may heat all the surrounding air to a temperature C with an infinitely long tube whose surface

is kept at a constant value C.

Show that any function $u(x, y)$ that is bounded and harmonic outside a finite closed interval approaches some limit as $x^2 + y^2 \longrightarrow \infty$.

2. The temperature on the surface of a sphere with centre at O and radius R is held at a constant value C. The initial temperature of the surrounding air is everywhere 0. Then, the temperature, $u(t, x, y, z)$ of the air at the instant t at the point (x, y, z) approaches the function (32.2) as $t \longrightarrow \infty$

Show by means of a transformation reversing the radius vectors that the solution to the external Dirichlet problem is unique in the class of bounded functions for the plane case and in the class of functions that approach zero as $OP \longrightarrow \infty$ for the three-dimensional case if the region G is infinite.

33. THE SECOND BOUNDARY-VALUE PROBLEM

1. The internal second boundary-value problem. Let us suppose that the region G in the x, y-plane is finite and bounded by a curve Γ of bounded curvature. As we stated earlier (Section 27), the second boundary-value problem consists in finding a function $u(x, y)$ that is harmonic in G, that is continuous in $G \cup \Gamma$, and whose derivative in the direction of the outer normal is equal at every point of the boundary of G to the value at that point of a given function f. We shall assume that the function f is continuous. This problem is also called the internal second boundary-value problem to distinguish it from the external second boundary-value problem that we shall consider in subsection 3 of Section 33. In Section 28, we showed that all solutions of the internal second boundary-value problem with a given function f can differ one from another only by constant terms.

A necessary condition for the existence of a solution to the internal second boundary-value problem is that the integral of f along the boundary of the region G be equal to zero.

We shall prove the necessity of this condition, assuming that $u(x, y)$ has continuous second derivatives inside G and that $\frac{\partial u}{\partial x}$ and $\frac{\partial u}{\partial y}$ have a continuous extension onto the boundary G. In Section 35, we shall dispense with these restric-

tions. In the present section, we shall prove the existence of a solution to the second boundary-value problem if the necessary condition stated above is satisfied.

Suppose that $u(x, y)$ is a solution of the second boundary-value problem in the region G and that $\frac{\partial u}{\partial n}=f(s)$ on Γ. Consider the integral

$$\iint\limits_G \left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right) dx\, dy.$$

It is equal to zero since the function u is harmonic. If we transform this integral into a line integral around the boundary Γ of the region G, we obtain, from Ostrogradskii's formula,

$$\int\limits_\Gamma \frac{\partial u}{\partial n}\, ds=0 \quad \text{or} \quad \int\limits_\Gamma f(s)\, ds=0, \tag{33.1}$$

since, by hypothesis, $\frac{\partial u}{\partial n}=f(s)$ on the boundary of the region. If the region G is multiply connected and if its boundary consists of a finite number of closed curves, the integral in equation (33.1) must be taken over all these lines and the direction chosen as the positive direction around each of these curves must be chosen in such a way that the region G will lie to the left of the boundary.

In the three-dimensional case, these same considerations may be used. In just the same way, we would see that the integral over the boundary G of the values of $\frac{\partial u}{\partial n}$ that are given on this boundary must be equal to zero.

2. For a two-dimensional simply-connected region G, the internal second boundary-value problem is easily reduced to the internal Dirichlet problem by the following procedure. Let us suppose that the internal second boundary-value problem has a solution u and that this solution and its first derivatives have continuous extensions onto $\overline{G}$. Let us then construct a function v in $\overline{G}$ in such a way that the Cauchy-Riemann equations

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x} \tag{33.2}$$

are satisfied in G. The function v, with derivatives defined by these equations exists since the condition

$$\frac{\partial^2 v}{\partial y\,\partial x} - \frac{\partial^2 v}{\partial x\,\partial y} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is satisfied. It is defined by these equations up to a constant term. It is easy to show that at every point in G the derivative of u in some direction $\vec{l}$ is equal to the derivative of v in the direction obtained by rotating $\vec{l}$ 90° counterclockwise. In just the same way, we may show that the derivative of u on the boundary of G in the direction of the normal to the boundary is equal to the derivative of v along the tangent to the boundary. Therefore, when we fix the value of v at any boundary point A in the region, we see that

$$v(B) - v(A) = \int_A^B f(s)\,ds, \tag{33.3}$$

at every point B of the boundary G, where ds denotes an element of length of the boundary of G. Since the integral of $f(s)$ over the entire boundary of G is equal to zero, equation (33.3) defines v on the boundary of G as a function that is everywhere continuous and single-valued.

It is easy to see that if u is harmonic, the function v defined by equations (33.2) is also harmonic. Therefore, when we know the values of v on the boundary of G, we can determine uniquely the value of v anywhere within G. Thus, if we assume that a solution $u(x,y)$ to the internal boundary-value problem exists in a region G for a given function $f(s)$ and that this solution and its first derivatives have continuous extensions onto $G \cup \Gamma$, we can determine $u(x,y)$ up to a constant term from equations (33.2) by constructing the corresponding solution $v(x,y)$ of the Dirichlet problem.

In the case of a three-dimensional region, analogous constructions are impossible,

3. The external second boundary-value problem. Suppose that a bounded single-valued region G has a smooth boundary Γ. Suppose that the points not belonging to $G \cup \Gamma$ form a region H with the boundary Γ. Find a function that

is harmonic in H, that is continuous in $H \cup \Gamma$, and that has a derivative in the direction of the normal directed outwardly from H at each point of the boundary of H that is equal to the value at that point of the given function f. Here, we shall require also that the solution $u(P)$ of the external second boundary-value problem be bounded in the case of two independent variables and that it tend to zero as the point P tends to infinity in the case of three and more independent variables.

In the case of two independent variables, the external second boundary-value problem can be reduced to the internal second boundary-value problem by a transformation reversing the radius vectors. Here, it is very significant that, because of the conformality of this kind of transformation all angles are conserved. Therefore, the normal to the boundary of the original region becomes a line normal to the boundary of the new region. The boundary function for the internal second boundary-value problem that we get in this way is obtained in the following way for the two-dimensional case. Keeping the notations used in considering the external Dirichlet problem, we obtain

$$u^*(P^*) = u(P), \quad OP \cdot OP^* = R^2,$$

$$f^*(s^*) = \frac{\partial u^*}{\partial n^*} = \frac{\partial u}{\partial n} \frac{dn}{dn^*} = f(s) \frac{dn}{dn^*}.$$

Here, s and s^* denote respectively points of the boundaries of the original and the new regions; n and n^* denote the normals to their boundaries; $\frac{dn}{dn^*}$ denotes the coefficient of stretching at a point on the boundary in the direction of the normal. Since the coefficient of stretching at a given point is independent of the direction in the case of a conformal transformation, in calculating $\frac{dn}{dn^*}$ we may assume that the directions of n and n^* pass through the centre O of the transformation. Then,

$$\frac{dn}{dn^*} = \frac{d(OP)}{d(OP^*)} = -\frac{R^2}{(OP^*)^2}.$$

For the external second boundary-value problem to have a solution, it is necessary and sufficient that the internal second boundary-value problem corresponding to it have a solution. And, as will be shown in Section 35, for this it is

necessary and sufficient that

$$0=\int\limits_{L^*} f^*(s^*)\,ds^*=\int\limits_{L} f(s)\frac{dn}{dn^*}\frac{ds^*}{ds}\,ds=\int\limits_{L} f(s)\,ds.$$

Here, L^* denotes the curve onto which L is mapped under the reciprocal transformation. Because of the conformality of this transformation

$$\frac{dn}{dn^*}\cdot\frac{ds^*}{ds}=1.$$

Therefore, by reducing the external second boundary-value problem to the internal one and by using the isolated-singularity theorem, we see that in the case of two independent variables two solutions of the same external second boundary-value problem can differ one from the other only by a constant term and condition (33.1) is a necessary and sufficient condition for the existence of a solution to the external second boundary-value problem.

In the case of three independent variables, we cannot use a reciprocal transformation to reduce the external second boundary-value problem to the internal one since in this case $\frac{\partial u^*}{\partial n^*}$ on the boundary is given not only in terms of $\frac{\partial u}{\partial n}$ but also in terms of the values of the unknown function u itself on Γ.

In the case of three or more independent variables, it is easy to show that the solution of the external second boundary-value problem is unique in the class of functions that approach zero as the point P approaches infinity. (Here 'approaches zero' is understood in the sense that $|u(P)|<\varepsilon$ for arbitrary $\varepsilon>0$ if the distance from the point P to the coordinate origin is sufficiently great.) We shall assume that the boundary Γ of the region H is such that each point of the boundary maybe touched by a sphere belonging to the region H.

Suppose that $u(P)$ is a harmonic function that is continuous in $H\cup\Gamma$. Suppose also $\frac{\partial u}{\partial n}=0$ Γ and that $u(P)\rightarrow 0$ as $P\rightarrow\infty$. Let us show that $u\equiv 0$.

Consider the region bounded by Γ and a sphere of sufficiently great radius that $|u(P)|<\varepsilon$ on that sphere. Since

$\frac{\partial u}{\partial n}=0$ on the boundary Γ, it follows from Theorem 1 of Section 28 and from the extreme-value theorem on harmonic functions that the function $u(P)$ assumes its largest and smallest values on the surface of the sphere; that is, $|u(P)|<\varepsilon$ throughout the region in question. Since $\varepsilon>0$ can be chosen arbitrarily small, $u(P)=0$ at every point P of the region H, which completes the proof.

34. POTENTIAL THEORY

1. In the next two sections, we shall obtain the solution of the fundamental boundary-value problems for Laplace's equation and also Poisson's equation (see Section 1) by the method of integral equations. This method is based on the representation of the solutions in the form of integrals that are often encountered in mechanics and physics and that have borrowed from these sciences the name potentials. These potentials are constructed with the aid of special kinds of particular solutions that have a definite type of singularity at a variable point.

Suppose that a point electric charge q is placed at a certain point O of the space (x, y, z). Then, from a familiar law of physics, this charge creates an electric field whose intensity $\mathbf{E}$ at an arbitrary point Q distinct from O is equal to

$$\mathbf{E}=kq\frac{\mathbf{r}}{r^3}$$

or, in components,

$$E_x=kq\frac{x-a}{r^3}; \quad E_y=kq\frac{y-b}{r^3}; \quad E_z=kq\frac{z-c}{r^3}. \tag{34.1}$$

Here, a, b, and c are the coordinates of the point O; x, y, and z are the coordinates of Q; $\mathbf{r}=\overrightarrow{OQ}$, and $r=OQ$, the coefficient of proportionality k depends on the choice of units.

The right-hand members of (34.1) are equal with opposite sign to the partial derivatives of the function

$$u(Q)=kq\frac{1}{r}+\text{const} \tag{34.2}$$

with respect to x, y and z respectively. This function is

called the potential of the given electrostatic field. It is customary to take the arbitrary constant in the right member of (34.2) equal to zero so that $u(Q) \longrightarrow 0$ as Q is moved infinitely far away. Also, in mathematical writings, it is customary to take $k=1$ for simplicity. Thus, we shall assume that the point charge of magnitude q creates a potential

$$u(Q)=\frac{q}{r}=\frac{q}{\sqrt{(x-a)^2+(y-b)^2+(z-c)^2}}. \qquad (34.3)$$

Since the potentials caused by several point charges are additive, the potentials caused by continuous distributions of charges are found in the form of the limit of a sum, that is, in the form of an integral. In particular, if a charge is distributed over a surface S with a surface density $\omega(A)$ (where $A \in S$), the potential caused by this charge is equal to

$$u(Q)=\iint\limits_{S} \frac{\omega(A)}{r(A, Q)}\, dS_A. \qquad (34.4)$$

Here, $r(A, Q)$ is the distance from A to Q and A is a variable point of integration, which is emphasised by the subscript in the differential. If the charge is distributed throughout a volume V with volume density $\rho(A)$ where $(A \in V)$, the potential caused by this charge is equal to

$$u(Q)=\iiint\limits_{V} \frac{\rho(A)}{r(A, Q)}\, dV_A. \qquad (34.5)$$

The right member of (34.4) is called the potential of a simple lamina and the right member of (34.5) is called a volume potential. Necessary assumptions ensuring the existence of these integrals will be given later.

Suppose now that two charges q and $-q$ lie on an axis l (Fig. 12) at a distance $h>0$ from each other. Suppose that they both move toward a point O and that the direction from $-q$ to q always coincides with the positive half of this axis. Then, the potential at an arbitrary point other than O is the difference between two quantities that tend to become equal to each other. Therefore, the potential in question tends to zero. However, if in the process of motion the charge q changes in such a way that

$$qh=p=\text{const},$$

the limit of the potential will be equal to

$$u(Q)=\lim q\left(\frac{1}{r'}-\frac{1}{r''}\right)=p\frac{\partial\left(\frac{1}{r}\right)}{\partial l}=-\frac{p}{r^2}\frac{\partial r}{\partial l}=$$
$$=p\frac{\cos(\overrightarrow{OQ},\ l)}{r^2}. \tag{34.6}$$

The limiting position of the charges is called in physics a dipole; the quantity p is called the dipole moment; the axis l is called the dipole axis. With the aid of point charges, a dipole can be attained only approximately (two charges of great magnitude at a small distance from each other). In the study of electrostatic fields, it is convenient to use the field of a dipole because of its simplicity along with the field of a point charge.

Suppose now that we have an oriented surface S, that is, one for which one side is designated as the outer side and the other the inner side. Suppose that a dipole is distributed on S with dipole-moment density $\tau(A)$ (where $A\in S$). Suppose also that the direction of the dipole axis at every point A coincides with the direction of the outer normal to S at the point A. Then, the potential caused by this dipole will be equal to

$$u(Q)=\iint\limits_S\frac{\tau(A)\cos(\overrightarrow{AQ},\ n_A)}{[r(A,\ Q)]^2}dS_A, \tag{34.4a}$$

where n_A is the outer normal to S at A. This integral is called the potential of a double lamina since this distribution of the dipole can be attained approximately if we put two distributions of charges on S with densities $\frac{1}{h}\tau(A)$ and $-\frac{1}{h}\tau(A)$ at a distance (along the normal to S) of h from one another provided $h>0$ is sufficiently small.

The right sides of (34.3) and (34.6) are harmonic functions everywhere in space except at the point O. This can be seen by direct calculation: one needs simply to show that (34.3) is harmonic since then (34.6) in the neighbourhood of each point other than O will be obtained as the uniform limit of harmonic functions. From this it follows, under slight assumptions regarding the density, that the potentials of the simple and double lamina are every-

where harmonic outside S.

Problem. Find the potential of a simple lamina from a uniform charge distribution on the surface of a sphere. Find the volume potential caused by a charge uniformly distributed throughout the sphere.

2. Suppose that the charge distribution is constant throughout space with respect to z. Then, the electrostatic field will also be independent of z. In this case, it will be sufficient for us to examine the entire picture of the distribution of charges and potentials in any of the planes $z=\text{const}$. Suppose that x and y are the coordinates in this plane. Instead of the intensity resulting from a point charge, here we need to consider the intensity at a point $Q(x, y)$ resulting from a charge of constant linear density q that is uniformly distributed along a straight line $x=a, y=b$. We denote the point (a, b) with the letter O. It follows from symmetry considerations that if $Q \neq 0$, the desired intensity will be equal to

$$E = qf(r)\,\boldsymbol{r}, \tag{34.7}$$

where $\boldsymbol{r} = \overrightarrow{OQ}$ and $r = |\boldsymbol{r}|$. To calculate $f(r)$, we take the point $(0,0)$ as the point O and $(r,0)$ as the point Q. Then,

$$E_y = E_z = 0, \quad E_x = k\int_{-\infty}^{\infty} \frac{rq}{(r^2+z^2)^{3/2}}\,dz = \frac{kq}{r}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\varphi\, d\varphi = \frac{2kq}{r}$$

$$(z = r\tan\varphi).$$

Therefore, (34.7) yields

$$f(r) = \frac{2k}{r^2},$$

from which we get

$$E = \frac{2kq}{r^2}\,\boldsymbol{r}$$

and, consequently, for an arbitrary position of the point Q in the x, y -plane,

$$E_x = \frac{2kq}{r^2}(x-a), \quad E_y = \frac{2kq}{r^2}(y-b).$$

These quantities will be equal, but with a change of sign, to the partial derivatives of the function

$$u(Q) = 2kq \ln \frac{1}{r} + \text{const} \tag{34.8}$$

with respect to x and y. This function is called the logarithmic potential or simply the potential. In mathematical works, it is customary to take $2k = 1$ and $\text{const} = 0$. Thus, in the case of a plane field, a point charge creates in the plane a potential

$$u(Q) = q \ln \frac{1}{r(O, Q)} = q \ln \frac{1}{\sqrt{(x-a)^2 + (y-b)^2}}. \tag{34.9}$$

We note that this potential cannot be found from (34.3) by direct integration along the charged line since we would then have a divergent integral.

The potential in a plane cause by a dipole is determined from this formula

$$u(Q) = p \frac{\partial \ln \frac{1}{r}}{\partial l} = p \frac{\cos(\overrightarrow{OQ}, l)}{r} \tag{34.10}$$

by a procedure analogous to that shown in subsection 1.

The right members of (34.9) and (34.10) are harmonic functions everywhere in the plane except at O (see subsection 1). The level curves of these functions (the equipotential curves) have the form shown in Fig. 13 (for a point charge) and Fig. 14 (for a dipole).

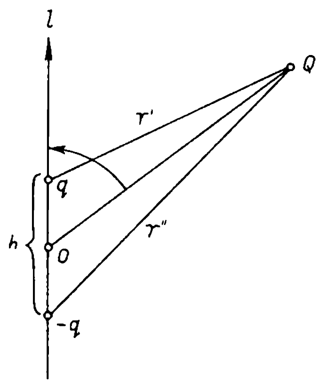


Fig. 12

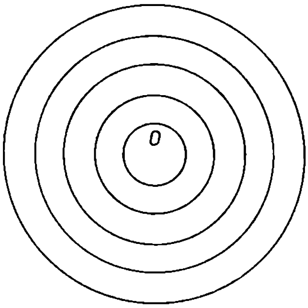


Fig. 13

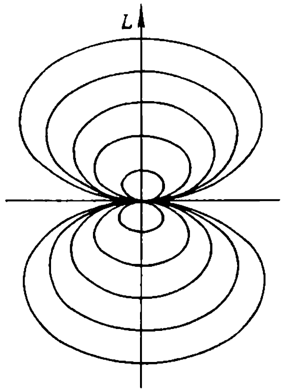


Fig. 14

The potentials due to a point charge and a dipole are expressed by (34.9) and (34.10) respectively. Instead of a volume potential, here we shall have a two-dimensional potential

$$u(Q) = \iint_G \rho(A) \ln \frac{1}{r(A, Q)} dS_A, \tag{34.11}$$

where G is a region in the plane. The potentials of a simple and of a double lamina for a plane have the following forms respectively:

$$u(Q) = \int_L \omega(A) \ln \frac{1}{r(A, Q)} dl_A, \tag{34.12}$$

$$u(Q) = \int_L \tau(A) \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} dl_A. \tag{34.13}$$

Here, L is a curve on the plane and n_A is a vector directed along the normal to L at the point A. We shall assume the curve L to be oriented, that is, one side of it is considered the outer side and the other the inner side. We shall assume the normal n_A to be outwardly directed.

In what follows, we shall treat only potential theory in a plane. This theory can be developed analogously in the space of an arbitrary number of dimensions.

Problem. Calculate the potential caused by a charge uniformly distributed on a circle. (The integral that would be obtained can be calculated with the aid of residue theory.)

3. In what follows, we shall consider a plane curve L with a continuously turning tangent with no points of self-intersection. Then, for an arbitrary point $P \in L$, we may place the coordinate axis in such a way that P will have the coordinates $x=0$, $y=0$ and a portion of L sufficiently close to P can be represented in the form

$$y=\varphi(x) \quad (-h \leqslant x \leqslant h;\ h>0), \tag{34.14}$$

Here, $\varphi'(x)$ exists and is continuous.

Suppose that a function $F(A, Q)$ is defined and continuous for $A \in L$ and Q varying arbitrarily on the plane but not coinciding with A and that F is not defined for $Q=A$. Then, the integral

$$w(Q)=\int_L F(A, Q)\, dl_A \tag{34.15}$$

is always defined and is a continuous function of Q when Q takes values not on L. The proof of this is elementary.

If $Q=P$ lies on L, the integral (34.15) becomes improper since the integrand is not defined for $A=P$. Then, we shall follow the usual practice of speaking of the convergence or divergence of the integral (34.15) depending on whether the limit

$$\lim_{\substack{A' \to P \\ A'' \to P}} \int_{L-l} F(A, P)\, dl_A, \tag{34.16}$$

exists or not. Here, l denotes the length of the arc of L with end points A' and A'' that holds the point P (see Fig. 15).

We shall say that the integral (34.15) converges uniformly at a point $P \in L$ if, for any $\varepsilon>0$, there exists a neigh-

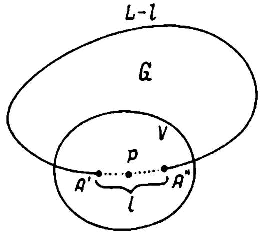


Fig. 15

bourhood V of the point P (see Fig. 15) and an arc l of the curve L containing the point P in its interior such that, for an arbitrary point $Q \in V$, the integral

$$\int_l F(A, Q)\, dl_A \tag{34.17}$$

converges to a number whose absolute value is less than ε. (The requirement of convergence is necessary only if Q lies on the part common to l and V.)

Theorem 1. Suppose that the integral (34.15) converges uniformly at some point $P \in L$. Then, for all points Q on L that are sufficiently close to P, the integral (34.15) converges and defines a function $w(Q)$ in some neighbourhood of the point P. This function is continuous at the point P.

Proof: Let us take an arbitrary $\varepsilon > 0$ and let us take a neighbourhood V and an arc l in accordance with the definition of uniform convergence at a point. Then, for an arbitrary point Q that is an interior point of the arc l and lies in V, the integral (34.17) converges. Therefore, the integral (34.15) converges for such points Q and the first assertion of the theorem (that $w(Q)$ is defined in some neighbourhood of the point P) is proven.

To show that $w(Q)$ is continuous at P, let us suppose that Q lies in V. Then,

$$\begin{aligned} |w(Q) - w(P)| &= \left| \int_L F(A, Q)\, dl_A - \int_L F(A, P)\, dl_A \right| \\ &\leqslant \left| \int_l F(A, Q)\, dl_A \right| + \left| \int_l F(A, P)\, dl_A \right| + \left| \int_{L-l} [F(A, Q) \right. \\ &\qquad \left. - F(A, P)]\, dl_A \right| \leqslant 2\varepsilon + \int_{L-l} |F(A, Q) - F(A, P)|\, dl_A. \end{aligned}$$

However, if l is fixed, this last integral will become less than ε if Q lies in a sufficiently small neighbourhood of the point P. This follows from the uniform continuity of the integrand as A varies on $L - l$ and Q varies in the neighbourhood of P referred to. Thus, if Q is sufficiently close to P,

$$|w(Q) - w(P)| < 3\varepsilon,$$

which, because of the arbitrariness of ε, proves the conti-

nuity of the function $w(Q)$ for $Q=P$. This completes the proof of Theorem 1.

Theorem 2. If $\omega(A)$ and $\tau(A)$ are continuous functions, the potentials of a simple and a double layer (34.12) and (34.13) are harmonic functions everywhere outside L.

Proof: The possibility of differentiating the functions (34.12) and (34.13) with respect to the coordinates of the point Q an arbitrary number of times if Q does not lie on L is proven in the same way that the possibility of differentiating a definite integral with respect to a parameter on which the integrand depends is proven in mathematical analysis. Therefore, the assertion in Theorem 2 follows immediately from the fact that the integrands in (34.12) and (34.13) are harmonic.

Theorem 3. The integral (34.12) converges when Q lies on L if $\omega(A)$ is a continuous function on L. Thus, the potential of a simple layer is a function defined on the entire plane. This function is continuous at every point on the plane.

Proof: From Theorems 1 and 2, it will be sufficient to prove the uniform convergence of the integral (34.12) at an arbitrary point $P \in L$. We take the point P for the coordinate origin and, directing the coordinate axes in a convenient direction, we write the equation for the portion of L close to P in the form (34.14). We denote this portion of L by l_h. Thus, we have

$$\left|\int_{l_h} \omega(A) \ln \frac{1}{r(A,\ Q)}\, dl_A\right| \leqslant \max_L |\omega(A)| \cdot \int_{l_h} |\ln r(A,\ Q)|\, dl_A$$
$$= \max_L |\omega(A)| \int_{-h}^{h} \left|\ln\sqrt{(x-a)^2+(y-b)^2}\right| \sqrt{1+[\varphi'(a)]^2}\, da, \tag{34.18}$$

where $b=\varphi(a)$.

If V and h are sufficiently small, the distance between an arbitrary point $Q(x,\ y)$ of the region V and an arbitrary point $A(a,\ b)$ on the line l_h will be less than 1, so that

$$0 \leqslant |x-a| \leqslant \sqrt{(x-a)^2+(y-b)^2} < 1,$$

and the estimate (34.18) then gives

$$\left|\int_{l_h} \omega(A) \ln \frac{1}{r(A, Q)} dl_A\right|$$

$$\leqslant \max_L |\omega(A)| \cdot \max_{l_h} \sqrt{1 + [\varphi'(a)]^2} \int_{-h}^{h} |\ln|x - a|| da$$

$$\leqslant \max_L |\omega(A)| \cdot \max_{l_h} \sqrt{1 + [\varphi'(a)]^2} \cdot 2 \int_0^{2h} |\ln a| da$$

if $Q \in V$. As is easily seen, the right member of this last inequality approaches 0 as h approaches 0 uniformly with respect to the point Q as Q varies in V. This completes the proof of Theorem 3.

Remark: We proved the convergence of the integral in the left member of (34.18) (for $Q \in l_h$) at the same time that we made the estimate of that integral, since an improper integral always converges if it converges absolutely.

In what follows, we shall denote by G the region bounded by a closed curve L with a continuously turning tangent and we shall denote by H the region consisting of those points that do not belong to $G \cup L$.

Theorem 4. The potential of a double lamina on L of unit density (that is, the integral (34.13) with $\tau(A) \equiv 1$) is equal to -2π when $Q \in G$; it converges and is equal to $-\pi$ when $Q \in L$; it is equal to zero when $Q \in H$.

Proof: Suppose that Q is an interior point of G and that A moves around L in the positive direction (see Fig. 16). We denote by α_{QA} the angle of inclination of the vector $\overrightarrow{QA}$ to the x-axis. Then, if we denote by $\overrightarrow{AB}$ the vector obtained by rotating $\overrightarrow{QA}$ 90° counterclockwise, we obtain*

$$\frac{d\alpha_{QA}}{dl} = \frac{\cos(\overrightarrow{AB}, \tau_A)}{r(A, Q)} = \frac{\cos(\overrightarrow{QA}, n_A)}{r(A, Q)} = -\frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)}.$$

Therefore,

$$\int_L \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} dl_A = -\int_L d\alpha_{QA} = -2\pi.$$

* This is easy to verify if we replace the differentials with increments and the arc Δl with the tangent to it at the point A.

The cases in which $Q \in L$ and $Q \in H$ are considered in an analogous manner. This completes the proof of Theorem 4.

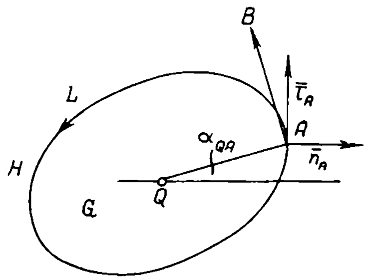


Fig. 16

We now turn to the general case.

Suppose that we are given the additional information that L is a plane closed curve with a continuously turning tangent and that it consists of a finite number of convex arcs and straight line segments. We shall call an arc convex if every straight line intersects it at not more than two points. Suppose that G is the region bounded by the curve L. Some of the arcs in L may be convex inwardly and others convex outwardly.

Theorem 5. The integral (34.13) converges when $Q \in L$ if $\tau(A)$ is a continuous function on L.

Hence, the potential of a double lamina $u(Q)$ is given by formula (34.13) everywhere on the plane. In general, it possesses a discontinuity of the first kind on L. More precisely, there is a continuous function $\underset{\smile}{u}(Q)$ in $G \cup L$ and a continuous function $\overset{\smile}{u}(Q)$ in $H \cup L$ such that

$$\left.\begin{aligned} u(Q) &= \underset{\smile}{u}(Q), \text{ when } Q \in G, \\ u(Q) &= \overset{\smile}{u}(Q), \text{ when } Q \in H, \\ u(Q) &= \frac{\overset{\smile}{u}(Q) + \underset{\smile}{u}(Q)}{2}, \text{ when } Q \in L, \\ \overset{\smile}{u}(Q) - \underset{\smile}{u}(Q) &= 2\pi\tau(Q), \text{ when } Q \in L. \end{aligned}\right\} \quad (34.19)$$

Proof: Let us take an arbitrary point $P \in L$ and let us consider the potential (34.13) and also the potential of the double lamina

$$u_1(Q) = \int_L \tau(P) \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} dl_A = \begin{cases} -2\pi\tau(P), \text{ when } Q \in G, \\ -\pi\tau(P), \text{ when } Q \in L, \\ 0, \text{ when } Q \in H. \end{cases}$$

We form the difference

$$u(Q) - u_1(Q) = \int_L [\tau(A) - \tau(P)] \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} dl_A \quad (34.20)$$

and show that the integral on the right converges uniformly to the point $Q = P$. It then follows on the basis of Theorem 1 that $u(Q)$ has a discontinuity at $Q = P$ of the same type as does $u_1(Q)$. This means that $u(Q)$ has discontinuities as $Q \longrightarrow P$ within G and as $Q \longrightarrow P$ within H. The value of $u(P)$ itself exists and is equal to the arithmetic mean of these two limiting values and the saltus of the function $u(Q)$ at P when P moves from G into H is equal to $2\pi\tau(P)$. The function $u(Q)$ considered for $Q \in G$ and extended onto L by assigning it its limiting values yields the function $u(Q)$, which is continuous in $G \cup L$. The situation with $Q \in H$ is analogous. This will be sufficient to prove Theorem 5.

To prove the uniform convergence of the integral (34.20) at the point P, let us take an arc l_h as in the proof of Theorem 3 and let us estimate an integral of the form (34.20) taken over l_h. We obtain

$$\left| \int_{l_h} [\tau(A) - \tau(P)] \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} dl_A \right| \leqslant \max_{l_h} |\tau(A) - \tau(P)| \int_{l_h} \left| \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} \right| dl_A.$$

We shall assume the arc l_h to be so small that it is constituted by not more than two convex arcs or straight line segments. It is easy to see that the expression

$$|\cos(\overrightarrow{AQ}, n_A)| \, dl_A$$

is equal to the projection of the element of arc dl_A onto the tangent at the point A to the circle of radius $r(A, Q)$ with centre at the point Q and that the expression

$$\left| \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} \right| dl_A.$$

is equal to the angle subtended at the point Q by the element dl_A.

Obviously, for any convex arc l that no ray originating at the point Q intersects in more than one point and also for an arbitrary straight line segment,

$$\int_l \left| \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} \right| dl_A \leqslant 2\pi.$$

Every convex arc l can be partitioned into two parts l_1 and l_2 each of which is intersected by no ray originating at the point Q at more than one point. Since the arc l_h consists of no more than four arcs (or line segments) possessing these properties, we have

$$\int_{l_h} \left| \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} \right| dl_A \leqslant 8\pi$$

and, consequently,

$$\max_{l_h} |\tau(A) - \tau(P)| \int_{l_h} \left| \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)} \right| dl_A \leqslant \max_{l_h} |\tau(A) - \tau(P)| \cdot 8\pi.$$

If we let h approach zero, then, because of the continuity of $\tau(A)$ the expression

$$\max_{l_h} |\tau(A) - \tau(P)| \cdot 8\pi$$

converges uniformly to zero for all Q. This completes the proof of Theorem 5.

Let us consider the normal derivative of the potential of a simple lamina. Suppose that $P \in L$ and that some function $F(Q)$ is defined in some neighbourhood of P. Then, we define

$$\frac{\partial F(P)}{\partial n^+} = \lim_{P' \to P} \frac{F(P') - F(P)}{r(P', P)},$$

$$\frac{\partial F(P)}{\partial n^-} = \lim_{P'' \to P} \frac{F(P) - F(P'')}{r(P, P'')}.$$

Here, n is the normal to the curve L drawn through the point P; n^+ is the distance along the portion of this normal outside G as measured from P, and n^- is the distance along that portion of this normal that lies inside G, again measured from P. We take as the positive direction of the nor-

mal the direction of n into the portion of the plane lying outside G. The point $P' \in H$ and the point $P'' \in G$.

We shall assume that L satisfies all the conditions enumerated just before Theorem 5 and, in addition, that its curvature is bounded. Then, we have

Theorem 6. The potential of a simple lamina $u(Q)$ given by formula (34.12) has at an arbitrary point $P \in L$ the derivatives $\frac{\partial u(P)}{\partial n^+}$ and $\frac{\partial u(P)}{\partial n^-}$. Also

$$\frac{\partial u(P)}{\partial n^+} = -\int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)} dl_A - \pi\omega(P), \qquad (34.21)$$

$$\frac{\partial u(P)}{\partial n^-} = -\int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)} dl_A + \pi\omega(P). \qquad (34.22)$$

The integrals on the right sides of (34.21) and (34.22) converge. It is assumed that $\omega(A)$ is a continuous function on L.

Proof: If Q lies on n_P but not on L, the derivative of $u(Q)$ in the direction of n_P exists and is determined by differentiating the integral (34.12) with respect to the parameter:

$$\frac{\partial u(Q)}{\partial n_P} = -\int_L \omega(A) \frac{\partial \ln r(A, Q)}{\partial n_P} dl_A$$

$$= \int_L \omega(A) \frac{\cos(\overrightarrow{AQ}, n_P)}{r(A, Q)} dl_A. \qquad (34.23)$$

Let us consider the potential of a double lamina $u_1(Q)$ caused by the distribution of a dipole over L with density $\omega(A)$. Then, if Q does not lie on L, we have

$$\frac{\partial u(Q)}{\partial n_P} + u_1(Q) = \int_L \omega(A) \frac{\cos(\overrightarrow{AQ}, n_A) - \cos(\overrightarrow{AQ}, n_P)}{r(A, Q)} dl_A. \qquad (34.24)$$

Let us show that the integral obtained converges uniformly at the point P if Q lies on n_P. Of course, we now need to change the definition of uniform convergence at the point P (see subsection 3); specifically, we need to require that the point Q lie not just anywhere in V but on the intersection

of n_P and the neighbourhood V of the point P. However, Theorem 1 is still conserved if we require that the point Q lie on n_P.

Suppose that l is a small arc of L close to the point P. Then, if $\max|\omega(A)|=C$, we have*

$$\left|\int_l \omega(A)\frac{\cos(\overrightarrow{AQ}, n_A)-\cos(\overrightarrow{AQ}, n_P)}{r(A, Q)}\,dl_A\right|$$

$$\leqslant C\int_l \frac{|\cos(\overrightarrow{AQ}, n_A)-\cos(\overrightarrow{AQ}, n_P)|}{r(A, Q)}\,dl_A$$

$$\leqslant 2C\int_l \frac{\left|\sin\dfrac{(\overrightarrow{AQ}, n_P)-(\overrightarrow{AQ}, n_A)}{2}\right|}{r(A, Q)}\,dl_A$$

$$=2C\int_l \frac{\left|\sin\dfrac{(n_A, n_P)}{2}\right|}{r(A, Q)}\,dl_A \leqslant C\int_l \frac{|(n_A, n_P)|}{r(A, Q)}\,dl_A \tag{34.25}$$

We assume that the curve L is of bounded curvature $\chi(A)$. Therefore,

$$|(n_A, n_P)|=\left|\int_{\breve{AP}}\chi(A)\,dl_A\right|<C_1|\breve{AP}|$$

and the left member of (34.25) will be no greater than

$$CC_1\int_l \frac{|\breve{AP}|}{r(A, Q)}\,dl_A. \tag{34.26}$$

If the arc l is sufficiently small, then for $A\neq P$

$$\frac{1}{\sqrt{2}}<|\sin(\overrightarrow{AP}, n_P)|<1$$

and

$$r(A, P)>\frac{1}{2}|\breve{AP}|.$$

* Here, we use that fact that

$$\cos\alpha-\cos\beta=2\sin\frac{\beta-\alpha}{2}\sin\frac{\alpha+\beta}{2} \text{ and } |\sin\alpha|\leqslant|\alpha|$$

for all α and β

Then, if we denote the projection of the point A onto n_p by A' (see Fig. 17), we have

$$r(A, Q) \geqslant r(A, A') > \frac{1}{\sqrt{2}} r(A, P) > \frac{1}{2\sqrt{2}} |\breve{AP}|$$

and the estimate (34.26) shows that the left member of (34.25) will be smaller than

$$CC_1 2\sqrt{2} \int_l dl = 2\sqrt{2} CC_1 |l|.$$

From this it is obvious that the left member of (34.25) approaches zero uniformly for all Q lying on n_p as $l \to 0$. Thus, the uniform convergence of the integral (34.24) is proven.

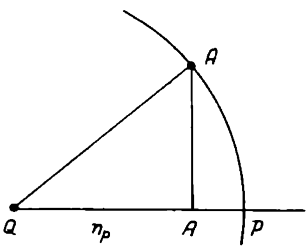


Fig. 17

The uniform convergence of the interval (34.24) at the point P implies, on the basis of Theorem 1 (suitably modified since Q lies on the intersection of V and n_p), that the integral is meaningful (that is, it converges) if $Q = P$ and that it has a limit as $Q \to P$ along the straight line n_p. This limit is equal to the value of the integral (34.24) when $Q = P$. In other words,

$$\lim_{P' \to P} \left[\frac{\partial u(P')}{\partial n_P} + u_1(P')\right] = \lim_{P'' \to P} \left[\frac{\partial u(P'')}{\partial n_P} + u_1(P'')\right]$$

$$= \int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_A) - \cos(\overrightarrow{AP}, n_P)}{r(A, P)} dl_A. \quad (34.27)$$

However, on the left side of (34.24), the nature of the discontinuity of the second term is determined by Theorem 5:

$$\lim_{P' \to P} u_1(P') = \hat{u}_1(P) = \int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)} dl_A + \pi\omega(P),$$

$$\lim_{P'' \to P} u_1(P'') = \underline{u}_1(P) = \int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)} dl_A - \pi\omega(P).$$

Therefore, it follows from (34.27) that the limits and the integral

$$\lim_{P' \to P} \frac{\partial u(P')}{\partial n_P}, \lim_{P'' \to P} \frac{\partial u(P'')}{\partial n_P}, \quad \int_L \omega(A) \frac{\cos(\overline{AP}, n_P)}{r(A, P)} dl_A$$

exist and that

$$\left.\begin{aligned} \lim_{P' \to P} \frac{\partial u(P')}{\partial n_P} &= -\int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)} dl_A - \pi\omega(P), \\ \lim_{P'' \to P} \frac{\partial u(P'')}{\partial n_P} &= -\int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)} dl_A + \pi\omega(P). \end{aligned}\right\} \quad (34.28)$$

With the aid of the theorem on finite increments, it is easy to show that if (1) a continuous function $f(x)$ is defined on some interval $[a, b]$, where $a < b$, (2) the derivative $f'(x)$ exists for $a < x < b$ and (3) the limit

$$\lim_{\substack{x \to a \\ (x > a)}} f'(x) \quad (34.29)$$

exists, then the derivative $f'(a)$ exists and is equal to (34.29). Of course, by $f'(a)$, we mean the right-hand derivative of $f(x)$, that is,

$$\lim_{\substack{\Delta x \to 0 \\ (\Delta x > 0)}} \frac{f(a + \Delta x) - f(a)}{\Delta x}.$$

Therefore, it follows from what was said above that $\frac{\partial u(P)}{\partial n^+}$ and $\frac{\partial u(P)}{\partial n^-}$ exist and that

$$\frac{\partial u(P)}{\partial n^+} = \lim_{P' \to P} \frac{\partial u(P')}{\partial n_P},$$

$$\frac{\partial u(P)}{\partial n^-} = \lim_{P'' \to P} \frac{\partial u(P'')}{\partial n_P}.$$

Formulae (34.21) and (34.22) follow from this and from (34.28). This completes the proof of Theorem 6.

Problem 1. Prove theorems analogous to Theorems 5 and 6 for the case in which L is not a closed curve but has a continuously turning tangent and consists of a finite number of convex arcs of bounded curvature.

2. Extend Theorems 3, 4, and 5 to the case in which G is a polygon.

Remark 1: All the theorems proven in the present section regarding the potentials of a simple or a double lamina remain valid if we assume only that the curve L is of bounded curvature.

2: All these theorems may naturally be carried over to the potentials of a simple or double lamina in three-dimensional space if we assume that the surface S over which the integrals corresponding to (34.4) and (34.4*) (for the potentials of a simple and double lamina respectively) are of bounded curvature. In this case, it turns out that the potential of a simple lamina is everywhere continuous but the potential of a double lamina and the normal derivatives of the potential of a single lamina around the point Q on a charged surface have respectively discontinuities of $4\pi\tau(Q)$, and $4\pi\omega(Q)$ instead of $2\pi\tau(Q)$ and $2\pi\omega(Q)$ in the case of a plane. Here, $\omega(Q)$ and $\tau(Q)$ denote respectively the density of the distribution of charges and dipoles on the surface S. It is assumed that these densities are continuous. In just the same way, all the reasoning of the present section can be carried over into three-dimensional space. The proof of these assertions can be found, for example, in the book by S.L. Sobolev, Uravneniya matematicheskoi fiziki (The equations of mathematical physics), Gostekhizdat, pp 208-228, 1954.

35. THE SOLUTION OF BOUNDARY-VALUE PROBLEMS USING POTENTIALS

1. The reduction of boundary-value problem for harmonic functions to integral equations. Suppose that L is a plane closed curve with a continuously turning tangent and continuous curvature and that it consists of a finite number of convex arcs and straight line segments†.

Suppose that a continuous function $f(P)$ is defined on L. Let us solve the internal Dirichlet problem, which consists,

†see footnote on page 298

as was shown in Section 27, in finding the function $u(Q)$ that is continuous in $G+L$ and harmonic in G and that satisfies the equation

$$u(P)=f(P). \tag{35.1}$$

on L.

We shall seek this harmonic function in the form of the potential of a double lamina (34.13) with unknown continuous density $\tau(A)$ of distribution of a dipole on L. On the basis of Theorems 2 and 5 of Section 34, this distribution is given by a function $\underset{\sim}{u}(Q)$ that is continuous in $G \cup L$ and harmonic for $Q \in G$. From (34.19), we have, for $P \in L$

$$\underset{\sim}{u}(P)=\int_L \tau(A)\,\frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)}\,dl_A - \pi\tau(P).$$

Therefore, for the boundary conditions (35.1), it is necessary and sufficient that the function $\tau(A)$ satisfy the Fredholm integral equation of the second kind

$$\tau(P)=\frac{1}{\pi}\int_L \tau(A)\,\frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)}\,dl_A - \frac{1}{\pi}f(P). \tag{35.2}$$

The external Dirichlet problem (see Section 32) is investigated in an analogous manner. If we seek a solution in the form of the potential of a double layer with unknown continuous density $\tau(A)$ of distribution of a dipole on L, we obtain an equation analogous to (35.2) for $\tau(A)$

$$\tau(P)=-\frac{1}{\pi}\int_L \tau(A)\,\frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)}\,dl_A + \frac{1}{\pi}f(P), \tag{35.3}$$

where $f(P)$ is a continuous function given on L.

† We shall assign to the curvature $X(A)$ at a point A on a curve L the sign determined by the positive direction of encirclement of L; that is

$$\chi(A)=\frac{d\alpha}{dl},$$

where α is the angle between the positive direction of the tangent and the x-axis. We shall consider the direction in which the point encircles L as positive if the region G remains on ones left.

The internal second boundary-value problem, as was pointed out in Section 27, consists in finding a function $u(Q)$ that is continuous in $G \cup L$ and harmonic in G and that possesses at each point L a derivative in the direction of the outer normal equal to the given continuous function $f(P)$.

Since we used $\frac{\partial u}{\partial n^-}$ in Section 34 to denote the derivative in the direction of the outer normal, we have for the solution $u(P)$ of the second boundary-value problem

$$\frac{\partial u(P)}{\partial n^-} = f(P) \quad (P \in L). \tag{35.4}$$

Let us seek a solution in the form of the potential of a simple layer (34.12) with an unknown function $\omega(A)$ which we shall assume to be continuous. On the basis of Theorem 6, Section 34, for the boundary condition (35.4) to be satisfied, it is necessary and sufficient that

$$\omega(P) = \frac{1}{\pi} \int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)} \, dl_A + \frac{1}{\pi} f(P). \tag{35.5}$$

The external second boundary-value problem is posed analogously. It reduces to the integral equation

$$\omega(P) = -\frac{1}{\pi} \int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)} \, dl_A + \frac{1}{\pi} f(P). \tag{35.6}$$

Remark: If we attempted to solve the internal Dirichlet problem by means of the potential of a simple layer with an unknown continuous density $\omega(A)$ of charge distribution, we would arrive at the equation

$$\int_L \omega(A) \ln \frac{1}{r(A, P)} \, dl_A = f(P) \qquad (P \in L). \tag{a}$$

This is a Fredholm integral equation of the first kind. The theory of such equations is much more complicated than the theory of equations of the second kind. As can be shown, equation (a) does not have a solution for all continuous $f(P)$. For example, if G is a circle of radius 1, equation (a) has no solution for $f(P) > 0$ since the left-hand side of (34.12) vanishes at the centre of this circle for an arbitrary function $\omega(A)$. For $f(P) > 0$ this is impossible because of the extreme-value theorem.

2. Investigation of the integral equations obtained. Let us set

$$K_1(P, A) = \frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)}; \quad K_2(P, A) = -\frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)}$$

$$(A \in L, \quad P \in L, \quad A \neq P).$$

Then,

$$K_1(A, P) = K_2(P, A).$$

Therefore, the kernels of equations (35.2) and (35.6) and also those of (35.3) and (35.5) are transposed.

The kernel $K_1(P, A)$ is defined and continuous when $A \in L$, $P \in L$, and $A \neq P$. However, for an arbitrary point $P_0 \in L$, the kernel $K_1(P, A)$ has a definite limit as $A \to P_0$ and $P \to P_0$ (where $A \neq P$). Suppose that T_A is the tangent to the curve L at the point A and that P_A is the projection of the point P onto T_A. Then, if the curvature $\chi(P_0)$ is positive, we have

$$P_0 \cos(\overrightarrow{AP}, n_A) = -|\sin(\overrightarrow{AP}, T_A)|.$$

Taking $|\sin \overrightarrow{AP}, T_A)|$ and $|\tan(\overrightarrow{AP}, T_A)|$, as equivalent and $r(A, P)$ and $r(A, P_A)$ as equivalent, we obtain

$$\lim \frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)} = -\lim \frac{|\tan(\overrightarrow{AP}, T_A)|}{r(A, P_A)}. \tag{35.7}$$

Let us take T_A as the x-axis and place the origin at the point A. We direct the y-axis into the interior of G. Then, the equation for the portion of L close to A is written in the form $y = \varphi(x)$. We denote by $\bar{x}$ the abscissa of the point P in this coordinate system. Then, by using Taylor's formula, we obtain

$$\frac{|\tan(\overrightarrow{AP}, T_A)|}{r(A, P_A)} = \frac{\varphi(\bar{x})}{\bar{x}^2} = \frac{1}{2}\varphi''(\theta\bar{x})$$

$$= \frac{1}{2}\chi(M)(1 + [\varphi'(\theta\bar{x})]^2)^{\frac{3}{2}}, \tag{35.8}$$

where the point M (with abscissa $\theta\bar{x}$) lies on L between A and P and where $\chi(M)$ is the curvature at the point M. It follows from (35.7) and (35.8) that

$$\lim \frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)} = -\frac{1}{2}\chi(P_0).$$

as $A \to P_0$ and $P \to P_0$ (for $A \neq P$). In the same way, we can show that this last equation is valid also in the case in which $\chi(P_0) \leqslant 0$.

If we extend the definition of the function $K_1(P, A)$ to the case in which $P = A$ by setting

$$K_1(A, A) = -\frac{1}{2}\chi(A),$$

the function obtained by doing so, which we shall also denote by $K_1(P, A)$, will be continuous over the set of variables for arbitrary $A \in L$ and $P \in L$ and hence is uniformly continuous. This applies also to $K_2(P, A)$.

We shall use the theory of integral equations with continuous kernel of the form

$$y(P) = \lambda \int_L K(P, A) y(A)\, dl_A + f(P),$$

which is expounded, for example, in my book on integral equations*.

Let us first show the following proposition, which we shall need in what we do afterwards.

Lemma 1. The potential of a simple layer

$$u(Q) = \int_L \omega(A) \ln \frac{1}{r(A, Q)}\, dl_A$$

approaches zero as the point Q is moved infinitely far away if and only if

$$\int_L \omega(A)\, dl_A = 0. \tag{35.9}$$

If condition (35.9) is not satisfied, the function $u(Q)$ increases without bound as Q is moved infinitely far away.

Proof: Let us take an arbitrary point O in the plane.

* PETROVSKII, I.G., *Lectures on the theory of integral equations*, Gostekhizdat, 50-54 (1951).

Then,

$$\int_L \omega(A) \ln \frac{1}{r(A, Q)} dl_A$$
$$= \int_L \omega(A) \ln \frac{1}{r(O, Q)} dl_A + \int_L \omega(A) \ln \frac{r(O, Q)}{r(A, Q)} dl_A$$
$$= \ln \frac{1}{r(O, Q)} \int_L \omega(A) dl_A + \int_L \omega(A) \ln \frac{r(O, Q)}{r(A, Q)} dl_A.$$

The second term in this sum approaches zero as Q is moved infinitely far away and the first term increases without bound in absolute value if and only if

$$\int_L \omega(A) dl_A \neq 0.$$

The assertion of the lemma follows.

Theorem 1. Equation (35.2) of the internal Dirichlet problem and equation (35.6) of the external second boundary-value problem have the same solution for a continuous function $f(P)$.

Proof: According to the first Fredholm theorem, we shall show that equations (35.2) and (35.6) have the same solution for an arbitrary continuous function $f(P)$ if we show that the homogeneous equations corresponding to them have only trivial solutions (that is, solutions that are identically equal to zero). Since equation (35.2) can be converted to equation (35.6), the second Fredholm theorem asserts that to prove Theorem 1, it will be sufficient to show that the homogeneous equation

$$\omega(P) = -\frac{1}{\pi} \int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)} dl_A \tag{35.10}$$

has only a trivial solution.

Suppose that $\omega(P)$ is a solution of (35.10). Let us show that

$$\int_L \omega(A) dl_A = 0.$$

Integrating the right and left members of equation (35.10)

over the curve L, we obtain

$$\int_L \omega(P)\,dl_P = -\frac{1}{\pi}\int_L \left[\int_L \omega(A)\frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)}\,dl_A\right] dl_P.$$

If we change the order of integration in the right member of this equation and apply Theorem 4 of Section 34, we obtain

$$\int_L \omega(P)\,dl_P = +\frac{1}{\pi}\int_L \omega(A)\left[\int_L \frac{\cos(\overrightarrow{PA}, n_P)}{r(A, P)}\,dl_P\right] dl_A$$
$$= \int_L \omega(A)\,dl_A,$$

that is,

$$\int_L \omega(P)\,dl_P = 0.$$

Consider the function

$$u(Q) = \int_L \omega(A)\ln\frac{1}{r(A, Q)}\,dl_A.$$

It follows from Lemma 1 of Section 35 that $u(Q)$ approaches zero as Q is moved infinitely far away. The function $u(Q)$ is harmonic outside L and since $\omega(P)$ satisfies equation (35.10), we have $\frac{\partial u}{\partial n^+} = 0$. But we showed in Section 33 that two solutions of the same external second boundary-value problem differ by a constant. Consequently, $u(Q) = \text{const}$ in H. Since $u(Q) \to 0$ as $Q \to \infty$, we have $u(Q) \equiv 0$ in H. It follows from the continuity of the potential of a simple layer that $u = 0$ on L. From the extreme-value theorem, $u \equiv 0$ in G, and, consequently $\frac{\partial u}{\partial n^-} = 0$. If we subtract equation (34.21) from (34.22), we obtain $\omega(P) \equiv 0$ since $\frac{\partial u}{\partial n^-} = 0$ and $\frac{\partial u}{\partial n^+} = 0$.

Theorem 2. The homogeneous equation

$$\omega(P) = \frac{1}{\pi}\int_L \omega(A)\frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)}\,dl_A, \qquad (35.11)$$

corresponding to equation (35.5) has only the linearly independent solution $\omega(P)$ and

$$u \int_L \bar{\omega}(A)\, dl_A \neq 0.$$

Proof: Let us show first that if a solution $\bar{\omega}(P)$ of equation (35.11) is not identically equal to zero, then

$$\int_L \bar{\omega}(A)\, dl_A \neq 0.$$

Consider the function

$$u(Q) = \int_L \bar{\omega}(A) \ln \frac{1}{r(A, Q)}\, dl_A.$$

The function $u(Q)$ is harmonic outside L. Since $\bar{\omega}$ satisfies equation (35.11), it follows from Theorem 6 of Section 34 that $\frac{\partial u}{\partial n^-} = 0$ on L. From Theorem 2 of subsection 3 of Section 28, we have $u \equiv \text{const}$ in $G \cup L$. If

$$\int \bar{\omega}\, dl_P = 0,$$

we know from Lemma 1 of Section 35 that $u(Q) \longrightarrow 0$ as Q is moved infinitely far away; that is, $u(Q)$ is a bounded solution of the external Dirichlet problem and is equal to a constant C on L. In Section 32, we showed the uniqueness of such a solution, so that $u(Q) = C$ in H. Since $u(Q) \longrightarrow 0$ as $Q \longrightarrow \infty$, it follows that $C = 0$, so that $u \equiv 0$ on the whole plane. It follows from Theorem 6 of Section 34 that $\bar{\omega}(P) \equiv 0$ on L.

The existence of at least one nontrivial solution $\bar{\omega}$ of equation (35.11) follows from the fact that the equation

$$\tau(P) = -\frac{1}{\pi} \int_L \tau(A) \frac{\cos(\overrightarrow{AP}, n_A)}{r(A, P)}\, dl_A,$$

which can be reduced to it, has the solution $\tau(P) \equiv \text{const}$ as can easily be checked.

Let us show that equation (35.11) cannot have two linearly independent solutions. Suppose that $\bar{\bar{\omega}}$ is some solution of (35.11) distinct from $\bar{\omega}$. A constant α can always be chosen so that

$$\int_L (\alpha\bar{\omega} + \bar{\bar{\omega}})\, dl_A = 0,$$

since

$$\int_L \bar{\omega}(A)\, dl_A \neq 0.$$

But we showed above that for a solution of (35.11) it follows from the equation

$$\int_L (\alpha\bar{\omega} + \bar{\bar{\omega}})\, dl_A = 0$$

that

$$\alpha\bar{\omega} + \bar{\bar{\omega}} \equiv 0.$$

This proves the theorem.

The function $\omega(P)$ has a simple physical interpretation. It is equal to the density of charge on L in the case in which $G \cup L$ is a conductor.

By using Theorem 2 and the third Fredholm theorem, we obtain

Theorem 3. Equation (35.3) of the external Dirichlet problem has a solution if and only if

$$\int_L f(A)\,\bar{\omega}(A)\, dl_A = 0. \tag{35.12}$$

When this condition is satisfied, the solution of equation (35.3) is determined up to an arbitrary constant term. Equation (35.5) of the internal second boundary-value problem has a solution if and only if

$$\int_L f(A)\, dl_A = 0. \tag{35.13}$$

When this condition is satisfied, the solution of equation (35.5) is determined up to a term of the form $C\bar{\omega}(P)$, where C is an arbitrary constant.

3. The solution of boundary-value problems. From Theorems 1 and 3 of the present section, we immediately

obtain results regarding the conditions under which boundary-value problems can be solved. First of all, it follows from Theorem 1 that under the restrictions that we have imposed, there always exists in G a unique solution of the internal Dirichlet problem represented in the form of the potential of a double layer. By virtue of the earlier proven uniqueness of the solution of the Dirichlet problem, we may say that the solution of the integral equation (35.2) is equivalent to the solution of the internal Dirichlet problem.

Furthermore, it follows from Theorem 3 that the solution of the internal second boundary-value problem exists for all functions $f(A)$ defined on the boundary such that

$$\int_L f(A)\, dl_A = 0.$$

Let us show that this condition is necessary for the internal second boundary-value problem with given function $f(A)$ to have a solution*. Suppose that $u(Q)$ is a function that is harmonic in G and continuous in $G \cup L$ and that $\frac{\partial u}{\partial n} = f(A)$ on L. Let us choose a constant C such that

$$\int_L (f(A) + C)\, dl_A = 0.$$

We proved above that there exists a function $v(Q)$ that is harmonic in G and continuous in $G \cup L$ such that

$$\frac{\partial v}{\partial n} = f(A) \cup C$$

on L.

The function $w = v - u$ is harmonic is G and continuous in $G \cup L$, and the derivative $\frac{\partial w}{\partial n} = C$ on L.

From Theorem 1 of subsection 3 of Section 28, we have $w =$ const and $C = 0$, since if w were not constant, $\frac{\partial \bar{w}}{\partial n}$ would have different signs at points of L where w attains its greatest and least values. From this it follows that

$$\int_L f(A)\, dl_A = 0.$$

* In Section 33, we showed that this condition is necessary under more stringent hypotheses.

It was shown in Section 28 that the solution of the internal second boundary-value problem is determined up to a constant term.

If we now turn to the solution of the external Dirichlet problem, we see on the basis of Theorem 3 that we cannot find a dipole distribution constituting a solution of this problem for an arbitrary boundary function. This is explained by the fact that, as is easily seen, every potential of a double layer (34.12) approaches zero at infinity and we proved the existence and uniqueness of a solution to the external Dirichlet problem in Section 32 by assuming only boundedness at infinity. In the case of a boundary function satisfying condition (35.12), there exists a solution of the external Dirichlet problem in the form of the potential of a double layer. For an arbitrary continuous function $f(P)$, we may proceed as follows: Let us define the function

$$f_1(P) = f(P) + C^*,$$

where we choose the constant C^* in such a way that $f_1(P)$ will satisfy equation (35.12). For this, we need to set

$$C^* = -\frac{\int\limits_L f(A)\,\bar{\omega}(A)\,dl_A}{\int\limits_L \bar{\omega}(A)\,dl_A}, \tag{35.14}$$

which we may do since, on the basis of Theorem 2,

$$\int\limits_L \bar{\omega}(A)\,dl_A \neq 0.$$

After we determine C^*, we solve equation (35.3) with f_2 substituted for f. Suppose that $\tau_1(P)$ is one of the solutions. Then, the function

$$u(Q) = \int\limits_L \tau_1(A)\,\frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)}\,dl_A - C^*.$$

will be a solution of the posed external Dirichlet problem. With regard to the constant term that appears in the dipole density given by equation (35.3), it does not show up in the solution of the external Dirichlet problem since the potential of a constant dipole distribution is equal to zero outside G

(see Theorem 4 of Section 34).

Finally, let us look at the external second boundary-value problem. As we have shown, the integral equation (35.6) corresponding to this problem can be solved for an arbitrary continuous function $f(P)$. Since the solution of the external second boundary-value problem is a function that is bounded at infinity, the potential of a simple layer with density defined as the solution of (35.6) will be the solution of the external second boundary-value problem if and only if it is bounded.

According to Lemma 1, for the potential of a simple layer to be bounded at infinity, it is necessary and sufficient that

$$\int_L \omega(A)\, dl_A = 0.$$

If we integrate equation (35.6), change the order of integration, and use Theorem 4 of Section 34, we obtain

$$\begin{aligned}\int_L f(P)\, dl_P &= \pi \int_L \omega(P)\, dl_P + \int_L \left[\int_L \omega(A) \frac{\cos(\overrightarrow{AP}, n_P)}{r(A, P)}\, dl_A\right] dl_P \\ &= \pi \int_L \omega(P)\, dl_P - \int_L \omega(A) \left[\int_L \frac{\cos(\overrightarrow{PA}, n_P)}{r(P, A)}\, dl_P\right] dl_A \\ &= 2\pi \int_L \omega(P)\, dl_P.\end{aligned}$$

Therefore, the condition

$$\int_L f(P)\, dl_P = 0$$

is necessary and sufficient for the potential (constructed with the use of equation (35.6)) of a simple layer to be bounded at infinity. According to Lemma 1, when this condition is satisfied the potential thus constructed must necessarily approach zero at infinity. Condition (35.13) is also a necessary condition for the external second boundary-value problem to have a solution. This follows from the necessity of condition (35.13) for the solvability of the internal second boundary-value problem and equation (33.4). Further, it follows from subsection 3 of Section 33 that the solution of the external second boundary-value problem is determined up to an arbitrary constant term.

4. The solution of boundary-value problems for a circle. If G is a circle, the solution of integral equations (35.2), (35.3), (35.5) and (35.6) is especially easy. Specifically, if we denote the radius of the circle by R, it is easily verified that for $A \in L$ and $P \in L$, we have

$$\cos(\overrightarrow{AP}, \boldsymbol{n}_A) = -\cos(\overrightarrow{AP}, \boldsymbol{n}_P) = -\frac{1}{2}\frac{r(A, P)}{R}$$

and equations (35.2) and (35.3) become

$$\tau(P) = \mp \frac{1}{2\pi R} \int_L \tau(A)\, dl_A \mp \frac{1}{\pi} f(P) \quad (P \in L), \tag{35.15$_{1,2}$}$$

and equations (35.5) and (35.6) become

$$\omega(P) = \pm \frac{1}{2\pi R} \int_L \omega(A) dl_A + \frac{1}{\pi} f(P) \quad (P \in L). \tag{35.16$_{1,2}$}$$

Let us solve $(35.15)_1$. We first set

$$\int_L \tau(A)\, dl_A = C$$

and integrate both sides of $(35.15)_1$ over L. We then obtain

$$C = -C - \frac{1}{\pi}\int_L f(P)\, dl_P; \quad C = -\frac{1}{2\pi}\int_L f(P) dl_P.$$

When we substitute this value of C into $(35.15)_1$, we obtain

$$\tau(P) = \frac{1}{4\pi^2 R} \int_L f(A)\, dl_A - \frac{1}{\pi} f(P) \quad (P \in L).$$

It now follows from (34.13) that for $Q \in G$, we have, on the basis of Theorem 4 of Section 34,

$$\begin{aligned} \underset{\sim}{u}(Q) &= \int_L \left[\frac{1}{4\pi^2 R} \int_L f(A)\, dl_A - \frac{1}{\pi} f(A)\right] \frac{\cos(\overrightarrow{AQ}, \boldsymbol{n}_A)}{r(A, Q)}\, dl_A \\ &= -\frac{1}{2\pi R} \int_L f(A)\, dl_A - \frac{1}{\pi} \int_L f(A) \frac{\cos(\overrightarrow{AQ}, \boldsymbol{n}_A)}{r(A, Q)}\, dl_A \\ &= \frac{1}{\pi} \int_L \left[\frac{\cos(\overrightarrow{QA}, \boldsymbol{n}_A)}{r(A, Q)} - \frac{1}{2R}\right] f(A)\, dl_A. \end{aligned}$$

Thus, we have obtained Poisson's integral, which we encountered in Section 29, in a different form.

Let us turn to equation $(35.15)_2$. The homogeneous equation corresponding to $(35.16)_1$ has the nontrivial solution $(P) \equiv \text{const} \neq 0$ (see Theorem 2). Thus, the conditions (35.12) and the equation

$$\int_L f(A)\, dl_A = 0$$

coincide here. If the condition

$$\int_L f(A)\, dl_A = 0$$

is satisfied, equation $(35.15)_2$ will have the solution

$$\tau(P) = \frac{1}{\pi} f(P) + C \quad (P \in L),$$

where C is arbitrary. In the general case, on the other hand, we have (see (35.14))

$$C^* = -\frac{1}{2\pi R} \int_L f(A)\, dl_A; \quad f_1(P) = f(P) - \frac{1}{2\pi R} \int_L f(A)\, dl_A;$$

$$\tau_1(P) = \frac{1}{\pi} f(P) - \frac{1}{2\pi^2 R} \int_L f(A)\, dl_A + C;$$

$$\tilde{u}(Q) = \int_L \left[\frac{1}{\pi} f(A) - \frac{1}{2\pi^2 R} \int_L f(A)\, dl_A\right] \frac{\cos(\overrightarrow{AQ}, n_A)}{r(A, Q)}\, dl_A + \frac{1}{2\pi R} \int_L f(A)\, dl_A$$

$$= -\frac{1}{\pi} \int_L \left[\frac{\cos(\overrightarrow{QA}, n_A)}{r(A, Q)} - \frac{1}{2R}\right] f(A)\, dl_A \quad (Q \in H).$$

Problem. Solve equations $(35.16)_{1,2}$ and obtain in connection with this the solution to the internal and external boundary-value problems for the circle. In solving the latter problem, use the formula

$$\int_0^\pi \ln(1 + p^2 - 2p\cos\varphi)\, d\varphi = 0, \quad -1 < p < 1.$$

5. Consider Poisson's equation

$$\Delta u = f(x, y). \tag{35.17}$$

We assume that the function $f(x, y) = f(P)$ is given in a bounded region G and that f is bounded and has continuous first partial derivatives. Let us solve the internal Dirichlet problem for this equation*. It will be sufficient to find any solution of equation (35.17) that is continuous in $\bar{G} = G + L$ without paying attention to the boundary function. For if v is such a solution, and if we set

$$u = v + w,$$

where w is the solution of the internal Dirichlet problem for Laplace's equation with the boundary condition

$$w\big|_L = u\big|_L - v\big|_L,$$

we see that u is a solution of the original problem. This also shows that the question of the existence and uniqueness of the solution of the internal Dirichlet problem for equation (35.17) amounts to just this same question regarding Laplace's equation.

Let us show that the function

$$v(P) = -\frac{1}{2\pi} \iint\limits_G f(A) \ln \frac{1}{r(A, P)} \, ds_A \tag{35.18}$$

is a particular solution of equation (35.17) (the logarithmic potential with charge density $-\frac{1}{2\pi} f(A)$).

Let us show first of all that the integral (35.18) converges and represents a continuous function of P on the entire plane. In analogy with Section 34, for this it will be sufficient to show that the integral (35.18) converges uniformly at an arbitrary point $P_0 \in \bar{G}$. Here, the definition of uniform convergence must be modified in a natural manner.

For an arbitrary $\rho > 0$, let us denote by $D_\rho(P_0)$ the interior of the circle with centre at P_0 and with radius ρ and let us denote by $G_\rho(P_0)$ the intersection of $D_\rho(P_0)$ and G. It

* That is, let us seek a solution of equation (35.17) that is continuous in $\bar{G}$ and that assumes on the boundary of $\bar{G}$ the values of a continuous function that is defined there.

is easy to show that, for arbitrary $\varepsilon > 0$, there exists a $\rho > 0$ such that, for an arbitrary point $P \in D_\rho(P_0)$, the integral

$$\iint\limits_{G_\rho(P_0)} f(A) \ln \frac{1}{r(A, P)}\, ds_A$$

converges and that its absolute value is less than ε. To show this, we denote by M the least upper bound of $|f|$ in G and we change to polar coordinates with poles at P. Then, if $\rho \leqslant \frac{1}{2}$ we have

$$\left| \iint\limits_{G_\rho(P_0)} f(A) \ln \frac{1}{r(A, P)}\, ds_A \right| \leqslant M \iint\limits_{G_\rho(P_0)} |\ln r(A, P)|\, ds_A$$

$$\leqslant M \int_0^{2\pi}\int_0^{2\rho} (-\ln r)\, r\, dr\, d\varphi = 4\pi M \rho^2 \left(\ln \frac{1}{2\rho} + \frac{1}{2}\right) \tag{35.19}$$

The last expression approaches zero uniformly as $\rho \to 0$ for all P belonging to $D_\rho(P_0)$.

Let us show that the integral (35.18) has continuous first partial derivatives. Let us note by (x, y) the coordinates of the point P and by (a, b) those of A. When we differentiate this integral formally with respect to x without worrying about convergence, we obtain

$$\varphi(P) = \frac{1}{2\pi} \iint\limits_G f(A)\, \frac{x-a}{[r(A, P)]^2}\, ds_A. \tag{35.20}$$

Just as with (35.19), it is easy to show that this interval converges uniformly at every point in $\overline{G} = G + L$ and that hence it represents the function itself, which is continuous on the entire plane. To show that $\varphi(P) \equiv v_x(P)$, we take an arbitrary fixed point* $P(x, y)$ and a point $P_1(x+h, y)$ (for $h \neq 0$). Then,

$$\left| \varphi(P) - \frac{v(P_1) - v(P)}{h} \right| = \frac{1}{2\pi} \left| \iint\limits_G f(A)\, \frac{x-a}{[r(A, P)]^2}\, ds_A \right.$$

cont. on next page

* By a procedure like that used in the proof of Theorem 2 (Section 34), one can easily show that the function $v(P)$ is harmonic outside the region G.

$$-\frac{1}{h}\left(-\iint_G f(A)\ln\frac{1}{r(A,P_1)}ds_A+\iint_G f(A)\ln\frac{1}{r(A,P)}ds_A\right)\Bigg|$$

$$\leqslant\frac{1}{2\pi}\left|\iint_{G_\rho(P)} f(A)\frac{x-a}{[r(A,P)]^2}ds_A\right|+\frac{1}{2\pi h}\left|\iint_{G_\rho(P)} f(A)\ln\frac{r(A,P)}{r(A,P_1)}ds_A\right|$$

$$+\frac{1}{2\pi}\left|\iint_{G-G_\rho(P)} f(A)\left(\frac{x-a}{[r(A,P)]^2}-\frac{1}{h}\ln\frac{r(A,P_1)}{r(A,P)}\right)ds_A\right| \qquad (35.21)$$

The first of the integrals on the right approaches zero as ρ approaches zero because of the uniform convergence of the integral (35.20) at the point P. The second also approaches zero if $0<|h|<\rho$. To prove this, we partition $G_\rho(P)$ into two parts: $G'_\rho(P)$, where $r(A,P)>r(A,P_1)$, and $G''_\rho(P)$, where $r(A,P)\leqslant r(A,P_1)$, and we note that $\ln(1+\delta)<\delta$ for $\delta>0$. We obtain

$$\frac{1}{2\pi h}\left|\iint_{G_\rho(P)} f(A)\ln\frac{r(A,P)}{r(A,P_1)}ds_A\right|$$

$$\leqslant\frac{M}{2\pi h}\left(\iint_{G'_\rho(P)}\ln\frac{r(A,P)}{r(A,P_1)}ds_A+\iint_{G''_\rho(P)}\ln\frac{r(A,P_1)}{r(A,P)}ds_A\right)$$

$$\leqslant\frac{M}{2\pi h}\left(\iint_{G'_\rho(P)}\frac{r(A,P)-r(A,P_1)}{r(A,P_1)}ds_A\right.$$

$$\left.+\iint_{G''_\rho(P)}\frac{r(A,P_1)-r(A,P)}{r(A,P)}ds_A\right)$$

$$\leqslant\frac{M}{2\pi}\left(\iint_{G'_\rho(P)}\frac{ds_A}{r(A,P_1)}+\iint_{G''_\rho(P)}\frac{ds_A}{r(A,P)}\right)^*$$

$$\leqslant\frac{M}{2\pi}\cdot 2\int_0^{2\pi}\int_0^{2\rho}\frac{1}{r}\,r\,dr\,d\varphi=4M\rho.$$

We may choose $\rho>0$ sufficiently small that the first and second integrals of the right-hand side of (35.21) will be

* Since $|r(A,P)-r(A,P_1)|\leqslant h$

less than $\epsilon/3$, where $\epsilon > 0$ is an arbitrary prestated number. When we fix this ρ we can, because of the fact that $|h|$ is decreasing, make the last integral in (35.21) less than $\epsilon/3$ since the integrand converges uniformly to zero in $G - G_\rho(P)$ as $|h| \to 0$. We treat v_y' in an analogous fashion:

$$v_x'(P) = \frac{1}{2\pi} \iint_G f(A) \frac{x-a}{[r(A, P)]^2} ds_A,$$

$$v_y'(P) = \frac{1}{2\pi} \iint_G f(A) \frac{y-b}{[r(A, P)]^2} ds_A.$$

Up to now, we have used only the boundedness of the continuous function $f(P)$. In what follows, we shall use the fact that the first partial derivatives of $f(P)$ are continuous. Let us fix the point $P_0 \in G$ and let us choose ρ sufficiently small that $D_\rho(P_0) \in G$. Then, the integral

$$v_1(P) = -\frac{1}{2\pi} \iint_{G - D_\rho(P_0)} f(A) \ln \frac{1}{r(A, P)} ds_A$$

has continuous partial derivatives of all orders in $D_\rho(P_0)$ and satisfies the equation

$$\Delta v_1 = 0, \tag{35.22}$$

since this integral can be differentiated under the integral sign with respect to the coordinates of P, which lies in $D_\rho(P_0)$, without any restrictions. This means that it will be sufficient to examine the integral

$$v_2(P) = -\frac{1}{2\pi} \iint_{D_\rho(P_0)} f(A) \ln \frac{1}{r(A, P)} ds_A.$$

Let us integrate the expression for $\frac{\partial v_2}{\partial x}$ by parts:

$$\frac{\partial v_2(P)}{\partial x} = \frac{1}{2\pi} \iint_{D_\rho(P_0)} f(A) \frac{x-a}{[r(A,P)]^2} ds_A$$

$$= \frac{1}{2\pi} \iint_{D_\rho(P_0)} \left\{ -\frac{\partial}{\partial a} [f(A) \ln r(A, P)] + f_a'(A) \ln r(A, P) \right\} da\, db$$

$$= -\frac{1}{2\pi} \int\limits_{C_\rho (P_0)} f(A) \ln r(A, P)\, db$$

$$- \frac{1}{2\pi} \iint\limits_{D_\rho (P_0)} f_a'(A) \ln \frac{1}{r(A, P)} ds_A, \qquad (35.23)$$

where $C_\rho(P_0)$ is the circle $D_\rho(P_0)$ and the integration over $C_\rho(P_0)$ is taken in the positive direction. It follows from what was proven above that the last integral has continuous first partial derivatives that are arbitrarily small for $P \in D_\rho(P_0)$, for sufficiently small ρ. The first integral on the right side of (35.23) can be differentiated in $D_\rho(P_0)$ with respect to x and y without any restrictions since the point P does not lie on the curve of integration. We can treat the expression for $\frac{\partial v_2}{\partial y}$ analogously. Thus, the existence and continuity of the second partial derivatives of $v_2(P)$ in $D_\rho(P_0)$, and hence those of $v(P)$ in G, are proven.

Furthermore, for $P \in D_\rho(P_0)$

$$\frac{\partial^2 v_2}{\partial x^2} = -\frac{1}{2\pi} \int\limits_{C_\rho (P_0)} f(A) \frac{x-a}{[r(A, P)]^2} db + \eta_1(P, \rho),$$

where $\eta_1(P, \rho)$ converges uniformly to zero as $\rho \to 0$. Analogously,

$$\frac{\partial^2 v_2}{\partial y^2} = \frac{1}{2\pi} \int\limits_{C_\rho (P_0)} f(A) \frac{y-b}{[r(A, P)]^2} da + \eta_2(P, \rho).$$

Let us change to polar coordinates with origin at $P_0(x_0, y_0)$. Then, on the basis of (35.22),

$$\Delta v = \frac{\partial^2 v_2(P_0)}{\partial x^2} + \frac{\partial^2 v_2(P_0)}{\partial y^2}$$

$$= \frac{1}{2\pi} \int\limits_{C_\rho (P_0)} \frac{f(A)}{[r(A, P)]^2} [(y_0 - b)\, da - (x_0 - a)\, db]$$

$$+ \eta_1(P_0, \rho) + \eta_2(P_0, \rho)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} f(x_0 + \rho \cos\varphi, y_0 + \rho \sin\varphi)(\sin^2\varphi + \cos^2\varphi)\, d\varphi + \eta_1 + \eta_2.$$

The last expression converges to $f(P_0)$ as $\rho \to 0$; thus,

$$\Delta v = f(P_0).$$

We note that the conditions imposed on the right side of equation (35.17) can be weakened somewhat. However, we cannot confine ourselves to requiring continuity and boundedness of the function $f(P)$ in G since then the integral (35.18) may not have second partial derivatives. In connection with this, I.I. Privalov has introduced the concept of a generalised Laplacian operator defined by

$$\Delta^*\varphi(P) = \lim_{\rho \to 0} \frac{4}{\rho^2}\left[\frac{1}{\pi\rho^2}\iint\limits_{D_\rho(P)} \varphi(A)\, ds_A - \varphi(P)\right].$$

It can be shown that if $\varphi(P)$ has continuous second partial derivatives in G, then for $P \in G, \Delta^* v(P)$ exists and is identically equal to $\Delta\varphi(P)$. At the same time, if $f(P)$ is continuous and bounded in a bounded region G and $v(P)$ is given by formula (35.18), then $\Delta^* v(P)$ exists and

$$\Delta^* v(P) \equiv f(P).$$

Remark: All the reasoning of this section can be carried over in a natural manner to the Newtonian potential (34.5) of a charged region in three-dimensional space. If we assume that the charge density $\rho(A)$ is continuous and bounded and that it has continuous first derivatives, then the potential itself $u(Q)$ is everywhere continuous. It is harmonic outside the charged region and satisfied Poisson's equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = -4\pi\rho$$

inside the charged region.

36. THE NETS METHOD FOR OBTAINING AN APPROXIMATE SOLUTION OF THE DIRICHLET PROBLEM

Suppose that a continuous function f is given on the boundary of a finite region G. Let us suppose that there exists a function u that is harmonic within G and that takes the given values of f on the boundary of G. To get an approximation of u, L.A.Lyusternik ‡ proposed the following method in 1925. For simplicity, we shall explain this method only for two-dimensional regions though it is equally applicable

‡ See p. 317 for footnote.

to regions of a greater number of dimensions. In our present exposition, we shall not include all the proofs at first. Unproven statements will be proven later in the text.

In the xy-plane, containing the region G, let us draw two families (a net) of straight lines parallel to the coordinates axes

$$x = mh \text{ and } y = nh,$$

where h is some positive number and m and n assume positive integral values, so that the entire region G is covered by squares of side h. The corners of these squares we shall call nodes or nodal points of the net. Our purpose is to determine the approximate values of u at the nodal points that lie within G. We shall denote these approximate values by u_h.

For any $\varepsilon > 0$, let us denote by Γ_ϵ the set of squares that have one corner at a distance no greater than ε from G. At each corner belonging to any of the squares Γ_ϵ, we set u_h equal to the value of f at the point on the boundary of G that is closest to this corner or, if there are several such points, the value at any one of them. For sufficiently small h and ε, the values of u_h that are defined in this way at the nodal points of Γ_ϵ would differ by as slight amount as we choose from the values of u at these points. This is true because the function that is equal to u within G and equal to f on the boundary of G is uniformly continuous in $\bar{G}$. Therefore, its values at the two points P_1 and P_2 belonging to $\bar{G}$ will be arbitrarily close together if the distance P_1P_2 is sufficiently small. In what follows, we shall assume that $h < \varepsilon$.

The points of the region G that lie within and on the boundaries of the squares that do not belong to Γ_ϵ form one

‡ See *Uspekhi matem. nauk,* 8, 115-124 (1941). Lyusternik did not assume the existence of a solution to the Dirichlet problem. He proved the existence of a solution to this problem under certain assumptions regarding the boundary G by the method of nets. However, his proof could not be extended directly to regions of more than two dimensions.

The existence of a solution to the Dirichlet problem for Laplace's equation with an arbitrary number of independent variables was proven by the method of nets for a broad class of regions in an article by PETROVSKII, I.G., *Uspekhi matem. nauk,* 8, 161-170 (1941).

or several polygons M. The nodal points lying on the boundary of each such polygon belong to Γ_ϵ and, therefore, the values of u_h are already defined at these points. The values of u_h at the nodal points lying within these polygons are defined as the solution of some system of linear equations the number of which is equal to the number of the values of u_h that are as yet undefined, that is, equal to the number of nodal points lying within G and not belonging to Γ_ϵ. This system of equations is set up in the following manner: For each nodal point (x, y), we write the equation

$$u_h(x, y) = \frac{u_h(x+h, y) + u_h(x-h, y) + u_h(x, y+h) + u_h(x, y-h)}{}$$

or

$$u_h(x+h, y) + u_h(x-h, y) + u_h(x, y+h) + u_h(x, y-h) - 4u_h(x, y) = 0 \qquad (36.1)$$

If any of the points $(x+h, y)$, $(x-h, y)$, $(x, y+h)$, or $(x, y-h)$ belongs to Γ_ϵ, the corresponding u_h is replaced in equation (36.1) by the value of u_h that was determined above at this point. It can be shown that the system of equations (36.1) always has a unique solution (Theorem 1) and that if we choose to begin with a sufficiently small ε and then decrease h sufficiently, we shall arrive at values of u_h that differ as slightly as we wish from the values of the functions $u(x, y)$ at the corresponding points (Theorem 2). The necessary degree of smallness of h depends on ε. By using the ordinary rules of algebra, the system of equations (36.1) would be difficult to solve if h is small and, consequently, the number of equations is high. However, for an approximate solution of the system (36.1), we can quite simply use the method of successive approximations (Theorem 3).

Equation (36.1) is the analogue in finite differences of Laplace's differential equation. Let us suppose that the function u has continuous derivatives of the first four orders in the region G in question. Let us suppose that the points

$$(x+h, y), \quad (x-h, y), \quad (x, y+h),$$

and $(x, y-h)$ and also the straight line segments between the point (x, y) and the points

$$(x+h, y), \ (x-h, y), \ (x, y+h), \ (x, y-h)$$

lie inside G. Then,

$$u(x+h,\ y)=u(x,\ y)+hu'_x(x,\ y) + \frac{h^2}{2}u''_{xx}(x,\ y)+\frac{h^3}{6}u'''_{xxx}(x,\ y)+\frac{h^4}{24}u''''_{xxxx}(\tilde{x},\ y),$$

$$u(x-h,\ y)=u(x,\ y)-hu'_x(x,\ y) + \frac{h^2}{2}u''_{xx}(x,\ y)-\frac{h^3}{6}u'''_{xxx}(x,\ y)+\frac{h^4}{24}u''''_{xxxx}(\tilde{\tilde{x}},\ y),$$

$$u(x,\ y+h)=u(x,\ y)+hu'_y(x,\ y) + \frac{h^2}{2}u''_{yy}(x,\ y)+\frac{h^3}{6}u'''_{yyy}(x,\ y)+\frac{h^4}{24}u''''_{yyyy}(x,\ \tilde{y}),$$

$$u(x,\ y-h)=u(x,\ y)-hu'_y(x,\ y)+\frac{h^2}{2}u''_{yy}(x,\ y) -\frac{h^3}{6}u'''_{yyy}(x,\ y)+\frac{h^4}{24}u''''_{yyyy}(x,\ \tilde{\tilde{y}}).$$

Here, $\tilde{x}$, $\tilde{\tilde{x}}$, $\tilde{y}$ and $\tilde{\tilde{y}}$ denote numbers lying respectively between x and $x+h$; x and $x-h$; y and $y+h$, and y and $y-h$. Obviously,

$$u(x+h,\ y)+u(x-h,\ y)+u(x,\ y+h)+u(x,\ y-h) - 4u(x,\ y)=h^2\left[u''_{xx}(x,\ y)+u''_{yy}(x,\ y)\right]+\frac{h^4}{6}M_1\theta,\ -1\leqslant\theta\leqslant 1,$$

where M_1 denotes the least upper bound of the values $|u''''_{xxxx}|$ and $|u''''_{yyyy}|$. Therefore, the left side of the equation

$$\frac{u(x+h,\ y)+u(x-h,\ y)+u(x,\ y+h)+u(x,\ y-h)-4u(x,\ y)}{h^2}=0$$

is equivalent to the expression $u''_{xx}+u''_{yy}$ up to a quantity of the order of h^2.

Theorem 1. The system of equations (36.1) always possesses a unique solution.

Proof: Let us rewrite this system in such a way that the left members of the equations will contain only the values of u_h at the interior nodal points of the polygons M and the right members will contain the values of u_h at the boundary nodal points of these polygons, that is, at the points Γ_ϵ. We recall that these last values have been defined. Therefore, we shall assume that the right sides of these equations

are known. Our system can then be thought of as being in the form

$$\sum_{j=1}^{N} a_{ij}u_j = f_i \quad (i = 1, 2, \ldots, N), \tag{36.2}$$

where N is the number of interior nodal points of the polygons M. We have numbered all the interior nodal points of the polygons and we have denoted the value of u_h at the j th point by u_j. The right members of (36.2) are linear combinations of the values of u_h at the boundary nodal points of the polygons M.

As we know from higher algebra, to show that the system (36.2) has a unique solution for every f_i it will be sufficient to show that the corresponding homogeneous system has only the zero solution. This last assertion can be proven by contradiction. Suppose that the system

$$\sum_{j=1}^{N} a_{ij}u_j = 0 \quad (i = 1, 2, \ldots, N) \tag{36.3}$$

has a nontrivial solution. We denote by B the largest of the numbers $|u_j|$ (for $(j = 1, \ldots, N)$, which, by hypothesis, is greater than zero. Without loss of generality, we may assume that B is equal to some one of the u_j since the case in which $-B$ is equal to some one of the u_j is reduced to the preceding case by changing the signs of all the u_j. Thus, we suppose that B is equal to some u_{j_0}. Since u_{j_0} is equal to the arithmetic mean of the values of u_j in the four neighbouring nodal points, u_j must be equal to B at each of these neighbouring nodal points (because if all the values of u_h at points adjacent to j_0 cannot be greater than B they also cannot be less than B). By neighbouring nodal points to the point (x, y), we mean the points

$$(x+h, \ y), \quad (x-h, \ y), \quad (x, \ y+h), (x, \ y-h).$$

Applying this reasoning to each of the nodal points adjacent to the j_0 th nodal point, we see that the u_j are also equal to B at these points. Continuing this reason, we see that $u_j = B$ at all the nodal points belonging to the boundary of some polygon M and adjacent to the same interior point P of that polygon. But this contradicts the fact that the right members f_i of all the equations (36.3) are equal to zero. This is true because the f_i are linear combinations of the

values of u_h at the points Γ_ε with coefficients equal to -1, as is easy to see by comparing (36.2) and (36.1). Therefore, the f_i cannot be zeros if u_h is equal to $B > 0$ at all points Γ_ε adjacent to P.

Theorem 2. If $u(x, y)$ is the exact solution of the Dirichlet problem, and u_h is the solution of the system (36.1) with the boundary conditions listed above, then, by choosing $\varepsilon > 0$ sufficiently small and then decreasing h sufficiently, we shall arrive at functions u_h that will differ by an arbitrarily small amount from the values of $u(x, y)$ at the corresponding points.

Proof: Let us show that $|u_h - u| < \delta$ at all nodes of the net for h sufficiently small. We place the coordinate origin inside the region G. We choose ε sufficiently small that $\max|u_h - u|$ will be less than $\delta/2$ at points of Γ_ϵ and we consider the auxiliary function v_h defined by

$$v_h = u_h - u - \frac{\delta}{2D^2}(D^2 - x^2 - y^2) - \frac{\delta}{2}.$$

Here, D denotes the diameter of the region G, that is, the least upper bound of distances between points belonging to it.

It is obvious that $v_h < 0$ at nodal points belonging to Γ_ϵ

$$\text{since } |u_h - u| < \frac{\delta}{2} \text{ and } D^2 > x^2 + y^2$$

at such points. Let us show that $v_h < 0$ and, consequently, $u_k - u < \delta$ at all nodal points belonging to M provided h is sufficiently small. We apply to the function v_h the operator Δ_h, which puts in correspondence with the function $\varphi(x, y)$ the function

$$\varphi(x+h, y) + \varphi(x-h, y) + \varphi(x, y+h) + \\ + \varphi(x, y-h) - 4\varphi(x, y).$$

Obviously,

$$\Delta_h v_h = \Delta_h u_h - \Delta_h u + \frac{\delta}{2D^2}\Delta_h(x^2 + y^2).$$

But

$$\Delta_h u_h = 0; \quad \Delta_h(x^2 + y^2) = 4h^2; \quad |\Delta_h u| < \frac{M_2 h^4}{6},$$

where M_2 is the least upper bound of the values of $|u''''_{xxxx}|$ and $|u''''_{yyyy}|$ in the polygons M. Therefore,

$$\Delta_h v_h > 4h^2 \frac{\delta}{2D^2} - \frac{M_2 h^4}{2} \tag{36.4}$$

and, for sufficiently small h,

$$\Delta_h v_h > 0. \tag{36.5}$$

But then it is easy to see that v_h cannot assume its largest value within any of the polygons M. And from this it follows that the quantity v_h is negative at all interior points of the polygons M since it is negative on the boundary of these polygons.

If we consider the function

$$w_h = u - u_h - \frac{\delta}{2D^2}(D^2 - x^2 - y^2) - \frac{\delta}{2},$$

we can see by a completely analogous process that $u - u_h < \delta$. If we compare the two results, we see that $|u - u_h| < \delta$ for sufficiently small h, which completes the proof.

Remark: If the exact solution $u(x, y)$ of the Dirichlet problem has bounded derivatives of the first four orders in G, the constant M_2 in the right member of (36.4) can be assumed to be independent of ε. Therefore, a sufficiently small h that will satisfy inequality (36.5) can also be chosen independently of ε. In view of this, we may simplify the construction shown above by taking as M the set of all squares with side h that, together with their boundaries, are contained in G.

Up to now, it has been completely immaterial to us in what order we number the nodal points lying inside the polygons M. Now, however, it will be important for us to take the following order.

The first nodal point must necessarily have as one of its neighbouring points a point lying on the boundary of one of these polygons. (We recall that we call any one of the points $(x+h, y)$, $(x-h, y)$, $(x, y+h)$, and $(x, y-h)$ a neighbouring point of the points (x, y).) The second nodal point must have as one of its neighbouring nodal points either a boundary point of one of the polygons M or the first point, etc. When we number the nodal points in this way, the system of equations (36.1) can be approximately solved as follows. Let us

first give arbitrary values to $u_1, u_2, \ldots, u_N$.

We shall denote these values by $u_k^{(0)}$ and we shall call them the zeroth approximation to the solution of the system (36.1). To describe the method of finding the successive approximations, it is convenient to imagine that all the $u_k^{(0)}$ are written down at the corresponding nodes of the net. Then, to obtain the next (first) approximation, we replace the value $u_1^{(0)}$ at the first nodal point with $u_1^{(1)}$ which is equal to the arithmetic mean of the values $u_k^{(0)}$ at the four nodal points neighbouring the first nodal point. Then, we replace the value $u_2^{(0)}$ at the second nodal point with the value $u_2^{(1)}$ equal to the arithmetic mean of the values written at the four nodal points neighbouring the second nodal point (one of which may be $u_1^{(1)}$), etc. When we treat all the interior nodal points in this way, we obtain the values of $u_k^{(1)}$ at them (for $k=1, \ldots, N$). The values of the second approximation $u_k^{(2)}$ (for $k=1, \ldots, N$) are obtained from the values of $u_k^{(1)}$ in the same way that these were obtained from the $u_k^{(0)}$. In an analogous way, we obtain $u_k^{(3)}, u_k^{(4)}, \ldots$ etc.

Theorem 3. For all k (for $k=1, 2, \ldots, N$)

$$u_k^{(1)} \longrightarrow u_k,$$

as $n \longrightarrow \infty$ where the u_k constitute the exact solution of the system (36.1).

Proof: Let us set

$$u_k^{(n)} - u_k = v_k^{(n)}.$$

Our purpose is to show that $v_k^{(n)} \longrightarrow 0$ (for $(k=1, 2, \ldots, N)$. as $n \to \infty$. We note first of all that the numbers $v_k^{(n+1)}$ are obtained from the $v_k^{(n)}$ (for $k=1, 2, \ldots, N$) in the same way as the $u_k^{(n+1)}$ are obtained from the $u_k^{(n)}$. Specifically, $v_k^{(n+1)}$ is the arithmetic mean of the values of $v_k^{(n)}$ at the four nodal points neighbouring the kth nodal point. If one of these neighbouring nodal points lines on the boundary of the polygon, $v_k^{(n)}$ is taken equal to zero. Therefore, if

$$\max \{ |v_1^{(0)}|, |v_2^{(0)}|, \ldots, |v_N^{(0)}| \} = A,$$

then

$$|v_1^{(1)}| \leqslant \frac{3}{4} A,$$

since one of the neighbouring points of the first nodal point is a boundary nodal point. Analogously, we obtain

$$|v_2^{(1)}| \leqslant \left(1 - \frac{1}{4^2}\right) A, \ |v_3^{(1)}| \leqslant \left(1 - \frac{1}{4^3}\right) A, \ \ldots,$$

$$|v_N^{(1)}| \leqslant \left(1 - \frac{1}{4N}\right) A = \alpha A,$$

while $\alpha < 1$.

Thus, in the same way, for all n and k, we obtain

$$|v_k^{(n)}| \leqslant \alpha^n A \quad \left(\alpha = 1 - \frac{1}{4N}\right), \tag{36.6}$$

from which it follows that

$$v_k^{(n)} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Theoretically, this relationship holds for an arbitrary choice of the zeroth approximation. However, in practice, if we take as our zeroth approximation numbers not too far from the exact solution of the Dirichlet problem, it will, as we might expect, be more convenient for a rapid calculation of suitable approximations to the exact solution of (36.1). The process of successive approximations usually stops at those values of n at which the values of $u_k^{(n)}$ cease to change appreciably upon increase in n. These $u_k^{(n)}$ are used as the approximate solution of the system (36.1).

If, for some h, we obtain several polygons M, the system (36.1) breaks into several independent systems each of which corresponds to one polygon. Each of these systems is solved independently of the others.

Our purpose in the above was to show that the successive approximations $u_k^{(n)}$ converge. The estimate (36.6) of the speed of convergence that we obtained in doing so is very crude. It would be possible to show that in actuality this process converges much more rapidly.

It should be noted that the successive approximations $u_k^{(n)}$ obtained by the simple method explained above converges only very slowly to the exact solution u_k of the system (36.1) when the number N of nodal points is large. There are various techniques that can be used to speed the convergence of these successive approximations to the exact solution and there are also other methods of obtaining an approximate solution of the system (36.1) which converge more rapidly and lead to the result more rapidly.

37. SUMMARY OF CERTAIN RESULTS FOR MORE GENERAL ELLIPTIC EQUATIONS

1. The Dirichlet problem for Laplace's equation in the case of two independent variables can be solved for an arbitrary simply connected region if the function given on the boundary is continuous. For a simply connected three-dimensional region, the Dirichlet problem does not always have a solution. In a three-dimensional region, any point P on the boundary will be regular if it is possible to touch the point in question from the outside with the vertex of a cone K obtained by rotating the curve

$$x_2 = f(x_1) = x_1^k,$$

where k is an arbitrary positive number, about the x_1-axis. This condition may be formulated more precisely as follows: In the space (x_1, x_2, x_3), containing the region G, it is possible to choose coordinate axes with origin at the point P in such a way that all points lying within the cone K and possessing positive abscissas x_1 not exceeding some positive number η lie outside the region G. On the other hand, Lebesque* and P.S. Urysohn** showed independently that a point P on the boundary of G will not be regular if it possesses a neighbourhood U_p such that for a suitable choice of coordinate axes all points in this neighbourhood not belonging to the region G lie within the cone generated by rotating the curve

$$x_2 = e^{-\frac{1}{x_1}}, \quad x_1 > 0.$$

about the x_1-axis. This will also be the case if we replace this curve with the curve

$$x_2 = F(x_1) = e^{-|\ln x_1|^{1+\varepsilon}} = x_1^{|\ln x_1|^{\varepsilon}},$$

where ε is any positive number.

For n-dimensional space (with $n > 3$) the function

$$\frac{x_1}{|\ln x_1|^{\frac{1}{n-3}}}, \tag{37.1}$$

* LEBESGUE, H., *Rendiconti del Circolo Matematico di Palermo*, **24**, 371-402 (1907).

URYSOHN, P.S., *Math. Zeitschrift*, **23, 155-158 (1925).

plays the role of the function $f(x_1)$ and the function

$$\frac{x_1}{|\ln x_1|^{\frac{1}{n-3}+\epsilon}}, \tag{37.2}$$

plays the role of the function $F(x_1)$, where ε is an arbitrary positive number. The equations of the corresponding cones are obtained by equating the expressions (37.1) or (37.2) to

$$\sqrt{x_2^2+\ldots+x_n^2}.$$

A necessary and sufficient condition for a point to be regular was found by Wiener*.

The question of the stability of the solution to the Dirichlet problem for Laplace's equation has been studied in connection with variation of the boundary of the region. Suppose that $G_1, G_2, \ldots, G_n, \ldots$ represent a sequence of regions that converge to the region G, that each of them contains the closed region $\bar{G}$, and that $\varphi(P)$ is an arbitrary function that is continuous throughout all space. We denote by $u_n(P)$ the function that is harmonic in G_n and that assumes on the boundary of G_n the values of $\varphi(P)$(for $n=1, 2, \ldots$).The Dirichlet problem is said to be stable within the region G if the sequence $\{u_n(P)\}$ converges as $n\to\infty$ at every point of G to the generalised solution (in the sense of subsection 3 of Section 31) to the Dirichlet problem corresponding to the boundary condition $u=\varphi(P)$ on the boundary of G.

A necessary and sufficient condition for stability of the Dirichlet problem within the region has been given by M.V. Keldysh and M.A. Lavrent'ev. An example has been exhibited of a simply-connected region in three-dimensional space for which the Dirichlet problem has a solution for an arbitrary continuous boundary function but is not stable within the region in question**.

2. Consider the linear elliptic equation with variable coefficients

$$\sum_{i,j=1}^{n} a_{ij}(x_1, \ldots, x_n)\frac{\partial^2 u}{\partial x_i \partial x_j}+\sum_{i=1}^{n} a_i(x_1, \ldots, x_n)\frac{\partial u}{\partial x_i}$$

cont. on next page

* See KELDYSH, M.V., *Uspekhi matem. nauk*, 8, 171-232 (1941).

** See the article by KELDYSH, 325.

$$+a(x_1, \ldots, x_n)u = f(x_1, \ldots, x_n), \tag{37.3}$$

where the quadratic form

$$\sum_{i,j=1}^{n} a_{ij}(x_1, \ldots, x_n)\alpha_i\alpha_j$$

is positive definite for arbitrary $x_1, \ldots, x_n$ belonging to the region in question. The sign of the coefficient $a(x_1, \ldots, x_n)$ has a significant effect on the possibility of solving the first boundary-value problem for this equation. If this coefficient assumes positive values, the first boundary-value problem for equation (37.3) may not have a solution or may have more than one solution if the region G is sufficiently great, even in the case of constant coefficients in equation (37.3). Thus, for example, the equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + 2k^2 u = 0 \tag{37.4}$$

has the solution $u_0 = \sin kx \sin ky$, which vanishes on the edges of a square Q with sides

$$x = 0;\ y = 0;\ x = \frac{\pi}{k};\ y = \frac{\pi}{k}.$$

On the other hand, it is easy to show that if equation (37.4) has a solution u_0 in a region G with a piecewise-smooth boundary Γ and if u_0 vanishes on Γ and possesses piecewise-continuous first derivatives in $G \cup \Gamma$, every other sufficiently smooth solution u of equation (37.4) must satisfy the relation

$$\int_{\Gamma} \frac{\partial u_0}{\partial n} u \, ds = 0 \tag{37.5}$$

on the boundary of the region. We shall obtain equation (37.5) if we integrate the left side of the equation

$$\iint_G u_0 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + 2k^2 u\right) dx\, dy = 0$$

by parts in such a way that the derivatives of u with respect to x and y will vanish in the integrals over the region G. Therefore, the first boundary-value problem for equation (37.4) when the region G is a square cannot have a smooth

solution if the function given on the boundary does not satisfy (37.5).

It can be shown that the first boundary-value problem for equation (37.3) either has a unique solution with an arbitrary continuous function defined on the boundary of the region G and for an arbitrary function f on the right side of the equation or has a solution only for those boundary functions and functions f on the right side that satisfy a finite number of conditions and the solution of the problem is unique.

In general, in solving the first boundary-value problem for the elliptic equation (37.3), the case in which the coefficient a is everywhere nonpositive is quite different from the case in which this coefficient assumes positive values at certain points. In the first case, the problem has a unique solution for an arbitrary continuous function given on the boundary of the region G if (1) the boundary of G is sufficiently regular and (2) the coefficients a_{ij}, a_i, and a and the function f satisfy a Hölder condition* in the region G**. However, if the coefficient a assumes positive values at certain points of G, to ensure the existence and uniqueness of the solution, we need to require also that the region G be sufficiently small. As V.V. Nemytskii has shown †, even for more general (nonlinear) equations, it is important here for the area of the region G to be sufficiently small, though its diameter may be arbitrarily large.

Oleinik has shown that, for an arbitrary region with $a \leqslant 0$ or for sufficiently small regions when a may be positive, the conditions that must be imposed on the boundary of the region for the Dirichlet problem to have a solution for this boundary and for every continuous function given on it do not depend on whether this problem is or is not solvable for Laplace's equation or equation (37.3)††.

* A function $\psi(x_1, \ldots, x_n)$ is said to satisfy a Hölder condition with exponent $\lambda > 0$ on a set M if there exists a constant K such that for any two points $(x_1, \ldots, x_n)$ and $(y_1, \ldots y_n)$ in the set M, the inequality

$$|\psi(x_1, \ldots, x_n) - \psi(y_1, \ldots, y_n)| \leqslant K\left[\sum_{i=1}^{n}(x_i - y_i)^2\right]^{\frac{\lambda}{2}}.$$

will be satisfied.

**MIRANDA, C., *Eguazioni alle derivate parziali di tipo ellititco*, Berlin, Springer (1957).

† NEMETSKII, V.V, *Matem. sbornik*, 41, 438-452 (1934).

††OLEINIK, O.A., *Matem. sbornik*, **24**, 1-14 (1949).

S.N. Bernstein proved the existence of a solution to the Dirichlet problem for a very broad class of nonlinear elliptic equations. The summary of these and other results concerning nonlinear elliptic equations is to be found in the journal Uspekhi matematicheskikh nauk, 8, 1941 (the article by Bernstein and Petrovskii, pp 8-26) and in the book by Miranda. This book expounds the most important divisions of the theory of linear and nonlinear second-order elliptic equations and contains a detailed bibliography.

3. The system of linear equations

$$\sum_{j=1}^{N} \sum_{0 \leq k_1+\ldots+k_n \leq n_j} A_{ij}^{(k_1\ldots k_n)}(x_1, \ldots, x_n) \frac{\partial^k u_j}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} = f_i(x_1, \ldots, x_n) \quad (i = 1, \ldots, N)$$

is said to be elliptic in a region G if the determinant

$$\left| \sum_{k_1+\ldots+k_n=n_j} A_{ij}^{(k_1\ldots k_n)}(x_1, \ldots, x_n) \alpha_1^{k_1} \ldots \alpha_n^{k_n} \right|$$

is nonzero for arbitrary real $\alpha_1, \ldots, \alpha_n$ not all zero and for arbitrary $x_1, \ldots, x_n$ in the region G. Ellipticity for a nonlinear system close to one of its solutions is defined analogously.

All sufficiently smooth solutions (that is, solutions possessing sufficiently many continuous derivatives) of elliptic equations and elliptic systems of equations are analytic if the left-hand members of these equations are analytic with respect to all their arguments. We assume that the right-hand members of these equations are all zero*. This was first shown by Bernstein in the case of second-order elliptic equations with two independent variables**.

4. If the homogeneous elliptic equation corresponding to (37.3) (that is, where $f \equiv 0$) with sufficiently smooth coefficients has in some bounded region G a unique solution to the first boundary-value problem for every continuous function given on the boundary of G, the theorem on a uniformly

* PETROVSKII, I.G., *Matem. sbornik*, 5, (47): 1, 3-70 (1939).

** BERNSTEIN, S.N., *Math. annalen*, 59, 20-76 (1904).

convergent sequence of solutions (analogous to Harnack's first theorem) is valid: specifically, if a sequence of solutions converges uniformly on the boundary of G, it also converges uniformly throughout the entire region G and it converges to a function satisfying equation (37.3).

5. We also have a theorem on monotonic sequences of solutions (analogous to Harnack's second theorem). Suppose that a bounded region G is such that, for every continuous function given on its boundary, the Dirichlet problem has one and only one solution. Then, if the sequence $u_n(x_1, \ldots, x_n)$ of solutions of the homogeneous form of equation (37.3) (i.e. $f \equiv 0$) converges at least one point of the region G and if

$$u_{n+1}(x_1, \ldots, x_n) \geqslant u_n(x_1, \ldots, x_n),$$

at all points in this region, the sequence $u_n(x_1, \ldots, x_n)$ converges uniformly in every region G' whose closure is contained in G.

6. If $a \equiv 0$ and $f \equiv 0$ in equation (37.3), then every solution of equation (37.3) has its largest and smallest values on the boundary of G. If $a \leqslant 0$ and $f \equiv 0$ in equation (37.3), then no nonconstant solution of equation (37.3) that is continuous in a closed region can have a greatest positive value or a smallest negative value in the interior of the region (cf. article by Oleinik cited in Section 28).

7. The solutions of the equation

$$\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) = 0 \tag{37.6}$$

possess the property of an arithmetic mean if we consider them in Riemann space with the corresponding metric. For such equations, we can construct solutions analogous to the potential of a point or a simple or double layer for Laplace's equation with two independent variables*. Analogous to these potentials are the so-called fundamental solutions that have

* FELLER, V., *Uspekhi matem. nauk*, 8, 232-248 (1941).

also been constructed for certain elliptic systems*.

8. Liouville's theorem for analytic functions can be carried over to certain second-order elliptic equations. Bernstein ** proved the following theorem. Every bounded solution of the equation

$$A(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy})\, u_{xx} + 2B(\ )\, u_{xy} + C(\ )\, u_{yy} = 0,$$

where A, B, C are bounded functions of their arguments and $AC - B^2 > 0$, that has continuous first and second partial derivatives on the entire plane, is a constant.

E.M. Landis investigated the behaviour of the solutions of linear second-order elliptic equations in various infinite regions. In particular, he proved theorems analogous to the Phragmén-Lindelöf theorem for analytic functions†.

9. If all functions u_i satisfying some homogeneous linear elliptic system of the form (3.2) with n independent variables and analytic coefficients and all their derivatives of the first $n_i - 1$ orders vanish simultaneously on some $(n-1)$-dimensional analytic surface, these functions are identically equal to zero throughout the entire region in which they satisfy the system in question.

This assertion can be obtained, for example, as a consequence of Holmgren's theorem on the uniqueness of the solution to the Cauchy problem for linear systems with analytic coefficients since elliptic systems do not possess real characteristics.

The uniqueness of the solution of the Cauchy problem is also proven for linear and quasilinear second-order elliptic equations with sufficiently smooth nonanalytic coefficients††.

* LEVI, E.E., *Uspekhi matem. nauk*, **8**, 249-292 (1941) and LOPATINSKII, Ya.B., *Ukrains. matem. zh.*, **3**, No. 1, 3-38 (1951).

BERNSTEIN, S.N., *Uspekhi matem. nauk*, **8, 75-81 (1941).

† LANDIS, E.M., *Dokl. Akad. Nauk SSSR*, **107**, No. 4, 508-511 (1956); *Uspekhi matem. nauk.* **14**: 1, 21-85 (85).

††LANDIS, E.M., *Dokl. Akad. Nauk SSSR*, **107**, No. 5, 640-643 (1956); CORDES, *Nachrichten Akad. Wiss. Göttingen*, No. 11, 239-258 (1956); LAVRENTEV, M.M., *Dokl. Akad. Nauk SSSR*, **112**, No. 2, 195-197 (1957). See also HEINZ, *Nachrichten Akad. Wiss. Göttingen*, No. 1, 1-12 (1955).

Furthermore, a number of results in this direction have been obtained for linear elliptic equations for higher orders and for linear elliptic systems*.

10. If all the coefficients a_{ij}, a_i, and a and the function f in the elliptic equation (37.3) have derivatives of the first $k-2$ orders (for $k>2$) that satisfy a Hölder condition in some finite region G, then all twice continuously differentiable solutions of this equation within G have derivatives of the first k orders satisfying a Hölder condition throughout every region G' whose closure is contained in G. An analogous assertion is also valid for nonlinear second-order elliptic equations.

The solution $u(x_1, \ldots, x_n)$ of the Dirichlet problem for equation (37.3) in the region G with boundary condition $u=\varphi$ on the boundary of G has derivatives of the first k orders in the closed region $\bar{G}$ that satisfy a Hölder condition provided (1) the boundary of the region G in a neighbourhood of each of its points can be represented by parametric equations $x_1=x_1(\xi_1, \ldots, \xi_{n-1}), \ldots; x_n=x_n(\xi_1, \ldots, \xi_{n-1})$, the right sides of which have kth order derivatives that satisfy a Hölder condition, (2) the boundary function φ has kth order derivatives satisfying a Hölder condition, and (3) the coefficients a_{ij}, a_i, and a and the function f have derivatives of the first $k-2$ orders in $\bar{G}$ that satisfy a Hölder condition.

Proofs of these results, discovered by Hopf and Schauder, are contained in the book by Miranda cited in footnote, p. 328

11. Vekua has investigated the question of the existence and uniqueness of a solution to the equation

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+a(x, y)\frac{\partial u}{\partial x}+b(x, y)\frac{\partial u}{\partial y} +c(x, y)u=f(x, y)$$

satisfying the condition

$$\alpha(x, y)\frac{\partial u}{\partial x}+\beta(x, y)\frac{\partial u}{\partial y}+\gamma(x, y)u=\varphi(x, y),$$

* CALDERON, *American Journal of Mathematics*, **80**, No. 1, 16-36 (1958); HÖRMANDER, *Mathematica Scandinavica*, 7, 177-190 (1959).

where a, b, c, f, α, β, γ, and φ are sufficiently smooth functions, on the boundary of the region G.

It turns out that the number of conditions that the functions f and φ must satisfy for this problem to have a solution and the number of independent solutions corresponding to the homogeneous problem ($f \equiv 0$, $\varphi \equiv 0$) depend on an integer n, called the index of the problem. The index of the problem n is equal to the increment in the argument of the function

$$\frac{\alpha(x, y) - i\beta(x, y)}{2\pi}$$

when the point (x, y) goes around the curve bounding G in the positive direction.

For simplicity in our formulations, let us suppose that the region G is simply connected and that $c \equiv 0, \gamma \equiv 0$. Then, if $n \geqslant 0$ our problem can be solved for arbitrary f and φ, and the number of linearly independent solutions of the corresponding homogeneous problem is $2n+2$ *. If $n < 0$, the problem has a solution only for those f and φ that satisfy certain conditions. The number of these conditions is $2n - 1$. In this case, the homogeneous problem has only a trivial solution.

Boundary-value problems of a more general kind have also been examined**.

12. The behaviour of solutions of the basic boundary-value problems for elliptic equations as a small parameter ε by the highest-order derivatives approaches zero has been studied. An asymptotic representation of these solutions has been obtained in the form of series in powers of ε. In a number of cases, this passage to the limit leads to a new boundary-value problem for an equation with $\varepsilon = 0$.

Boundary-value problems have also been studied for equations that are elliptic inside the region in question and parabolic on a portion of its boundary (or on some interior subset).

A detailed survey of the results concerning questions of this sort is to be found in the article by M.I. Vishik,

* Compare the conditions under which the first boundary-value problem for equation (37.3) in subsection 2 of this section has a solution.

** VEKUA, I.N., *New methods of solving elliptic equations*, Gostekhizdat, 1948; *Generalised analytic functions*, Fizmatgiz (1959).

A.D. Myshkis and O.A. Oleinik 'Differentsial'nye uravneniya s chastnymi proizvodnymi' (Partial differential equations) in the book Matematika v SSSR za sorok let (Mathematics in the USSR in the last forty years), Vol. I, Fizmatgiz, pp 599-603, 1959.

13. A continuous function $u(x_1, \ldots, x_n)$ satisfying the integral identity (see Section 9)

$$\int_G \cdots \int [uM(\sigma) - f\sigma]\, dx_1 \ldots dx_n = 0$$

for an arbitrary function $\sigma(x_1, \ldots, x_n)$, whose first m derivatives are continuous in G and which is equal to zero in a neighbourhood of the boundary G is called a generalised solution of the elliptic equation

$$\sum_{0 \leqslant k_1 + \ldots + k_n \leqslant m} A^{(k)}_{k_1 \ldots k_n}(x_1, \ldots, x_n) \frac{\partial^k u}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} = f(x_1, \ldots, x_n) \tag{37.7}$$

Here,

$$M(\sigma) \equiv \sum_{k=0}^{m} (-1)^k \sum_{k_1 + \ldots + k_n = k} \frac{\partial^k (A^{(k)}_{k_1 \ldots k_n} \sigma)}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}}.$$

It turns out that every generalised solution of the elliptic equation (37.7) in the region G possesses continuous derivatives of the first m orders and satisfied this equation in the ordinary sense if in the region G the function f has continuous derivatives of the first $2[n/2]$ orders and the coefficients $A^{(k)}_{k_1 \ldots k_n}$ of the kth derivatives in equation (37.7) have continuous derivatives of the first $k + 2[n/2]$ orders (for $k = 0, 1, \ldots, m$).

The smoother the coefficients in equation (37.7), the smoother will be the solution of the equation. In particular, if all the coefficients $A^{(k)}_{k_1 \ldots k_n}$ and the function f have continuous derivatives of every order in G, then every generalised solution of equation (37.7) will also have derivatives of every order.

A partial differential equation every generalised solution of which possesses derivatives of every order is said to be

hypoelliptic. Obviously, every linear elliptic equation with infinitely differentiable coefficients is hypoelliptic. An example of an equation that is hypoelliptic but not elliptic is the heat-flow equation (see Chapter IV). Sufficient (and in the case of equations with constant coefficients, necessary) conditions for hypoellipticity have been found by Hörmander*.

14. Just as the Dirichlet problem was a characteristic boundary-value problem for Laplace's equation, the problem of finding a solution u of the 'semi-harmonic' equation

$$\Delta^m u \equiv \left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\ldots+\frac{\partial^2}{\partial x_n^2}\right)^m u=0$$

inside some region G from its values and the values of its first $m-1$ normal derivatives on the boundary of G is a characteristic boundary-value problem for this equation. Certain important problems in the theory of elasticity lead to this problem with $m=2$ and $n=2$ or 3. The existence and uniqueness of an ordinary solution to this problem are proven under the assumption of sufficient smoothness of the boundary of the region G and of the functions given on it. For $m=2$ and $n=2$, it is sufficient to require that the region G be bounded by a finite number of closed curves the coordinates of each of which are three times continuously differentiable functions of the arc length and that the functions given on the curves and their first derivatives be continuous along the arc. S.L. Sobolev proved the existence and uniqueness of an ordinary solution of this problem under quite general assumptions concerning the boundary of G. He assumed that this boundary consists of several pieces of different dimensionality. It then turned out that on a piece of dimensionality $n-r$, the values of the function u and its first m - [r/2] - 1 derivatives must be given.

Sobolev's solution can be generalised in the sense that the function u and its derivatives do not necessarily assume the given values at all boundary points but only 'in mean'. (For a precise definition of 'in mean' as used here, see the book by Sobolev Nekororye primeneniya funktsional'nogo analiza v matematicheskoi fizike (Certain applications of functional

* HÖRMANDER, *Communications on pure and applied mathematics*, **11**, No. 1, 197-218 (1958).

analysis in mathematical physics), pp 111-113, 1950.)

15. For a certain class of linear elliptic systems known as strongly elliptic systems, Vishik* has investigated the question of the solvability of boundary-value problems analogous to the first and second boundary-value problems for a second-order elliptic equation. (In particular, this class includes a linear elliptic equation of the general form (37.7).) Just as in the case of equation (37.3), it turns out that either such a problem has a unique solution for arbitrary given boundary functions and the right members of the system or the solution is not unique and for the existence of a solution it is necessary that a finite number of conditions be satisfied for the boundary functions and right members. Sufficient conditions for the existence and uniqueness of a solution to the first and second boundary-value problems that the coefficients of the system must satisfy have been found. We note that just as in the case of equation (37.3), for strongly elliptic systems a solution to the first boundary-value problem always exists and is unique in sufficiently small neighbourhoods.

As is shown by certain examples exhibited by A.V. Bitsadze as far back as 1948, the first boundary-value problem with homogeneous boundary conditions for an elliptic system with two independent variables can have an infinite number of linearly independent solutions in an arbitrarily small circle**. In recent years, a number of interesting results have been obtained concerning the existence and uniqueness of solutions to the boundary-value problems for general elliptic systems with many independent variables†.

* VISHIK, M.I., *Matem. sbornik*, **29**, (71): 3, 615-676 (1951); *Dokl. Akad. Nauk SSSR*, **86**, No. 4, 645-648 (1952). See also GARDING, *Mathematica Scandinavica*, **1**, 55-72 (1953); BROWDER, *Annals of Math. Studies*, **33**, 15-51 (1954); GUSEVA, O.V., *Dokl. Akad. Nauk SSSR*, **102**, No. 6, 1069-1072 (1955); NIRENBERG, *Communications on pure and applied mathematics*, **8**, No. 4, 649-675 (1955).

** BITSADZE, A.V., *Uspekhi matem. nauk*, **3**:6 (28), 241-242 (1948).

† AGMON, DOUGLIS, NIRENBERG, *Communications on pure and applied mathematics*, **12**, No. 4, 623-727 (1959).

Parabolic equations

38. THE FIRST BOUNDARY-VALUE PROBLEM. THE EXTREME-VALUE THEOREM

1. As the simplest example of a parabolic equation, let us consider the heat-flow equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x_1^2} + \ldots + \frac{\partial^2 u}{\partial x_n^2}.$$

The basic properties of the solutions to this equation do not depend on n. For simplicity, we shall confine ourselves to a consideration of the case $n = 1$.

Parabolic equations are most frequently encountered in the study of the processes of heat flow and diffusion (see Section 1).

The following is a typical boundary-value problem for a

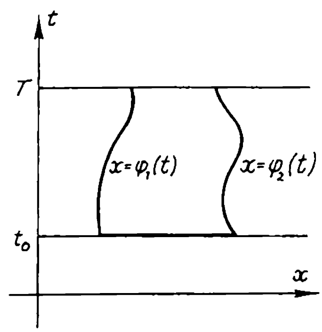


Fig. 18

parabolic equation. Let us denote by G the curvilinear quadrilateral in the tx-plane bounded by the line segments $t=t_0$ and $t=T$ (where $T>t_0$) and the curves $x=\varphi_1(t)$ and $x=\varphi_2(t)$, where φ_1 and φ_2 are continuous functions and $\varphi_2(t)>\varphi_1(t)$ for $t_0\leqslant t\leqslant T$ (see Fig. 18). We denote by Γ that portion of the region G consisting of the line segment $t=t_0$ and the two curves $x=\varphi_1(t)$ and $x=\varphi_2(t)$. (In Fig. 18, this portion of the boundary is indicated by the heavy lines.)

We need to find a function $u(t, x)$ that is continuous in the region G and on its boundary and that satisfies inside G the equation

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2} \tag{38.1}$$

and that assumes on Γ the values of a continuous function f given on Γ.

In the theory of equation (38.1), this problem plays the same fundamental role that the Dirichlet problem plays in the theory of Laplace's equation. It is called the first boundary-value problem for the heat-flow equation. In the case in which the region G is a rectangle $Q: 0<x<l,\ 0<t<T$, the problem, for example, of finding the temperature $u(t, x)$ in a thermally isolated rod when we know its initial temperature at $t=0$ and we know the temperature at the ends of the rod at all subsequent instants leads to the first boundary-value problem for the heat-flow equation. In solving this problem, it is significant that we are seeking a solution for $t>0$. As we shall show later, the analogous problem for negative values of t has in general no solution. In contrast with the equation for a vibrating string (20.1), equation (38.1) changes significantly when we replace t with $-t$. This is a typical equation of an irreversible process.

In what follows, we shall consider only continuous solutions of the heat-flow equation without always making specific mention of the point.

2. The extreme-value theorem and its corollaries.

Theorem. Every solution $u(t, x)$ of the heat-flow equation (38.1) that is defined and continuous in the curvilinear quadrilateral G and on its boundary assumes a largest and a smallest value on the boundary Γ, that is, either on the lower base of the curvilinear quadrilateral or on its lateral sides.

Since the theorem as applied to a minimum is reduced to the theorem as applied to a maximum by a change of sign of u, we shall confine ourselves to proving it only as regards to maximum.

The method of proof is analogous to the method of I.I. Privalov for proving the extreme-value theorem for harmonic functions†. We denote by M the maximum of the function $u(t, x)$ in $G \cup \Gamma$ and we denote by m the maximum of the values of $u(t, x)$ on Γ. Let us suppose that there exists a solution u for which $M > m$, that is, we assume that the theorem is not true for a maximum. Suppose that this function assumes the value M at a point (t^*, x^*), where $t^* > t_0$ and $\varphi_1(t^*) < x^* < \varphi_2(t^*)$.

Consider the function

$$v(t, x) = u(t, x) + \frac{M-m}{4l^2}(x - x^*)^2,$$

where l is equal to $\max\limits_{t_0 \leq t \leq T} \varphi_2(t) - \min\limits_{t_0 \leq t \leq T} \varphi_1(t)$.

On the lateral sides of G and on its lower base,

$$v(t, x) \leq m + \frac{M-m}{4} = \frac{M}{4} + \frac{3}{4}m = \theta M,$$

where $0 < \theta < 1$, and

$$v(t^*, x^*) = M.$$

Consequently, $v(t, x)$, like $u(t, x)$ does not assume a maximum value either on the lower base of G or on its lateral sides. Suppose that $v(t, x)$ assumes its maximum value at a point (t_1, x_1), where $t_1 > t_0$ and $\varphi_1(t_1) < x_1 < \varphi_2(t_1)$. At this point $\frac{\partial^2 v}{\partial x^2} \leq 0$ and $\frac{\partial v}{\partial t} \geq 0$. (If $t_1 < T$, then $\frac{\partial v}{\partial t}$ must be equal to 0 at that point. On the other hand, if $t_1 = T$, we have $\frac{\partial v}{\partial t} \geq 0$.) Therefore, at the point (t_1, x_1) we have

$$\frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} \geq 0 \tag{38.2}$$

On the other hand,

$$\frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} = \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - \frac{M-m}{2l^2} = -\frac{M-m}{2l^2} < 0,$$

† See *Matem. sbornik*, 32, 464-471 (1925).

which contradicts (38.2).

Corollaries. (1) The solution of the first boundary-value problem in a curvilinear quadrilateral G is unique. This is true because the difference between two solutions is equal to zero on the lower base and on the lateral sides of G and, on the basis of the extreme-value theorem, is identically equal to zero.

(2) The solution of the first boundary-value problem for the heat-flow equation depends continuously on the functions given on the lateral sides and on the lower base of the curvilinear quadrilateral G. This also follows from the fact that the difference between two solutions u_1 and u_2 of equation (38.1) in the curvilinear quadrilateral G has its greatest and lowest values on the lower base or on the lateral sides of G.

Problem 1. Prove the following strengthened extreme-value theorem: a nonconstant solution $u(t, x)$ of the heat-flow equation (38.1) that is defined and continuous in a closed curvilinear quadrilateral $\overline{G}$ cannot have its greatest or lowest value at any interior point of the upper base of $\overline{G}$.

2. Show that if a nonconstant function $u(t, x)$ that is continuous in the closed rectangle $\{t_0 \leqslant t \leqslant T, X_1 \leqslant x \leqslant X_2\}$ satisfies within this rectangle the heat-flow equation (38.1) and the inequality $u(t, x) \geqslant u(T, X_1)$ (or $u(t, x) \geqslant u(T, X_2)$ and if the derivative $\frac{\partial u}{\partial x}$ exists at the point (T, X_1) (resp. (T, X_2)), then $\frac{\partial u}{\partial x} > 0$ at (T, X_1) (resp. $\frac{\partial u}{\partial x} < 0$ at (T, X_2)).

3. Prove the uniqueness of a solution $u(t, x)$ to the heat-flow equation (38.1) that is continuous and has a continuous first partial derivative with respect to $\frac{\partial u}{\partial x}$ in the rectangle $\{t_0 \leqslant t \leqslant T, X_1 \leqslant x \leqslant X_2\}$ and that satisfies the following conditions:

$$u(0, x) = \varphi(x), \quad \frac{\partial u}{\partial x} - a_1(t)\, u = \varphi_1(t)$$

at $x = X_1$, and

$$\frac{\partial u}{\partial x} + a_2(t)\, u = \varphi_2(t)$$

at $x = X_2$, where a_1 and a_2 are given continuous nonnegative functions of t.

39. SOLUTION OF THE FIRST BOUNDARY-VALUE PROBLEM FOR A RECTANGLE BY THE FOURIER METHOD

For the rectangle Q:

$$0 \leqslant x \leqslant l, \quad 0 \leqslant t \leqslant T,$$

the first boundary-value problem can be worded as follows: Find a function $u(t, x)$ that is continuous in Q and that satisfies equation (38.1) in the interior of Q, the initial condition

$$u(0, x) = \varphi(x)$$

and the boundary conditions

$$u(t, 0) = f_1(t), \; u(t, l) = f_2(t), \qquad (0 \leqslant t \leqslant T). \tag{39.1}$$

on the boundary of Q. Here, it is assumed that the functions $\varphi(x)$, $f_1(t)$, and $f_2(t)$ are continuous and that $\varphi(0) = f_1(0)$, and $\varphi(l) = f_2(0)$.

Since equation (38.1) remains unchanged when we replace t with $t + t_0$ and x with $x + x_0$, everything that we shall say with regard to the rectangle Q will remain valid for every other rectangle with sides parallel to the x- and t-axes.

In this section, we shall prove the existence of a solution to the first boundary-value problem for the rectangle Q by the Fourier method.

The basic defect in this method consists in the fact that it is directly applicable only to a problem with the homogeneous boundary conditions $u(t, 0) = 0, u(t, l) = 0$ although the existence of a solution to the first boundary-value problem can be proven by another method for arbitrary $u(t, 0) = f_1(t)$ and $u(t, l) = f_2(t)$ that satisfy the condition stated above. However, if we know how to find any solution $v(t, x)$ at all to equation (38.1) that satisfies some definite boundary conditions $v(t, 0) = \tilde{f}_1(t)$ and $v(t, l) = \tilde{f}_2(t)$, we can apply the Fourier method to solving this equation for an arbitrary given initial condition $u(0, x) = \varphi(x)$ and for these boundary conditions $u(t, 0) = \tilde{f}_1(t)$ and $u(t, l) = \tilde{f}_2(t)$. To do this, all that we need to do is to find by the Fourier method a solution $u^*(t, x)$ of equation (38.1) satisfying the initial condition $u^*(0, x) = \varphi(x) - v(0, x)$ and the homogeneous

boundary conditions $u^*(t, 0)=0$ and $u^*(t, l)=0$. Once $u^*(t, x)$ is found, the function

$$u(t, x)=u^*(t, x)+v(t, x)$$

will be the solution to the original problem.

The existence of a solution to the first boundary-value problem in the general case will be proven by the method of nets in Section 42.

Let us seek a solution in Q to equation (38.1) that satisfies the following conditions: $u(0, x)=\varphi(x)$, where $\varphi(x)$ is a continuously differentiable function that vanishes at $x=0$ and $x=l$; $u(t, 0)=u(t, l)=0$ for $0 \leqslant t \leqslant T$.

In a manner analogous to what was done in solving the mixed problem for a hyperbolic equation, we shall seek a function $u(t, x)$ in the form of a series

$$u(t, x)=\sum_{k=1}^{\infty} T_k(t) X_k(x), \tag{39.2}$$

in which each term satisfies the heat-flow equation and vanishes at $x=0$ and $x=l$. The conditions must then be such that

$$\frac{T_k'(t)}{T_k(t)}=\frac{X_k''(x)}{X_k(x)}=-\lambda_k^2, \tag{39.3}$$

$$X_k(0)=X_k(l)=0. \tag{39.4}$$

Here, the $-\lambda_k^2$ denote certain constants. That they must be negative is proven just as in Section 20. From equations (39.3) and conditions (39.4), we find

$$X_k(x)=A_k \cos \lambda_k x+B_k \sin \lambda_k x; \quad A_k=0; \quad \lambda_k l=k\pi,$$

where $k=1, 2, \ldots$ Thus,

$$X_k(x)=B_k \sin \frac{k\pi}{l} x.$$

When we substitute the value of λ_k that we have found into equation (39.3), we obtain

$$T_k(t)=B_k^* e^{-\frac{k^2\pi^2}{l^2} t},$$

where the B_k^* are constants. Therefore,

$$u(t, x) = \sum_{k=1}^{\infty} C_k e^{-\frac{k^2\pi^2}{l^2}t} \sin \frac{k\pi}{l} x. \tag{39.5}$$

Here, the C_k are constants. They are determined from the initial condition

$$u(0, x) = \varphi(x) = \sum_{k=1}^{\infty} C_k \sin \frac{k\pi}{l} x. \tag{39.6}$$

Since we have assumed that $\varphi(x)$ is a continuously differentiable function that vanishes at $x = 0$ and $x = l$, the coefficients C_k can be determined from Fourier's formulas. They are bounded. The series (39.6) with these coefficients converges uniformly and absolutely to $\varphi(x)$, as we know from the theory of trigonometric series, since

$$0 < e^{-\frac{k^2\pi^2}{l^2}t} \leqslant 1,$$

for $t \geqslant 0$ the series (39.5) also converges absolutely and uniformly for $t \geqslant 0$. Therefore, the function $u(t, x)$ defined by this series is continuous in the rectangle $Q: 0 \leqslant x \leqslant l$; $0 \leqslant t \leqslant T$, and it assumes the given values on its lower base and its lateral sides. It remains to show that within Q and on its upper base, this function satisfies the heat-flow equation. To show this, it will be sufficient to show that the series obtained from (39.5) by differentiating termwise once with respect to t and that obtained by differentiating termwise twice with respect to x also converge absolutely and uniformly for $t \geqslant t_0 > 0$ for every $t_0 > 0$. But this last assertion follows from the fact that, for every positive t_0

$$\frac{k^2\pi^2}{l^2} e^{-\frac{k^2\pi^2}{l^2}t_0} < 1,$$

if k is sufficiently great.

In just the same way, we can prove that the function $u(t, x)$ has continuous derivatives of all orders with respect to x and t for $t > 0$ By using this fact, we can easily show that, keeping the original zero conditions at $x = 0$ and $x = l$, we cannot in general extend the function $u(t, x)$ that we have just constructed to negative values of t in such a way that

it will satisfy equation (38.1). For this, it is necessary that $\varphi(x)$ have derivatives of all orders. If such an extension were possible, we would obtain a solution to the heat-flow equation in some rectangle Q_1

$$0 < x < l, \quad -\varepsilon \leqslant t \leqslant 0,$$

that vanishes for $x = 0$ and $x = l$. If $\frac{\partial u}{\partial x}(-\varepsilon, x)$ is continuous for $0 \leqslant x \leqslant l$, we can, by applying the same reasoning to this rectangle that we did above for Q, show that the function $u(0, x)$, that is, $\varphi(x)$, must have derivatives of all orders. (It can be shown that the assumption that $\frac{\partial u}{\partial x}(-\varepsilon, x)$ is continuous at $x = 0$ and $x = l$ is not essential).

Even if the function $u(0, x) = \varphi(x)$ is such that the first boundary-value problem can be solved for it in the rectangle Q_1 under zero conditions at the ends of the interval $(0, l)$ and the initial condition $u(0, x) = \varphi(x)$, it will be possible to change this solution by an arbitrarily great amount for arbitrarily small negative values of t by changing the function $\varphi(x)$ and its derivatives of arbitrary fixed order k by an arbitrarily small amount. For this, as is easily shown, it is sufficient to add to the original solution any term with sufficiently great index in the series (39.5) with an arbitrarily small constant coefficient. Therefore, whereas the first boundary-value problem for the heat-flow equation is correctly stated for positive values of t, it is incorrectly stated for negative t if the initial conditions apply to $t = 0$ (see Section 8). Here, we see again the lack of symmetry between positive and negative values of t for the heat-flow equation (38.1).

Problem. Show that the solution $u(t, x)$ of the first boundary-value problem for the half-plane $\{0 \leqslant x \leqslant l,\ 0 \leqslant t < \infty\}$ under the conditions $u(0, x) = \varphi(x)$, $u(t, 0) = 0$ and $u(t, l) = 0$ approaches 0 uniformly with respect to x as $t \to \infty$

40. THE CAUCHY PROBLEM

1. Statement of the problem. Determine the function $u(t, x)$ that is continuous and bounded for $t \geqslant 0$, that satisfies the heat-flow equation (38.1) for $t > 0$, and whose values for $t = 0$ represent a given continuous bounded func-

tions $\varphi(x)$ defined for all real values of x.

The problem of the flow of heat in an infinitely long thermally insulated rod, for example, leads to this problem.

2. The extreme-value theorem for a strip and its corollaries. Every solution $u(t, x)$ of equation (38.1) that is continuous and bounded in the strip

$$S\{0 \leqslant t < T \leqslant \infty, -\infty < x < \infty\},$$

satisfies the inequalities

$$M \geqslant u(t, x) \geqslant m,$$

$$\text{where } M = \sup_{-\infty < x < \infty} u(0, x), \quad m = \inf_{-\infty < x < \infty} u(0, x).$$

everywhere in this strip.

Let us show that $u(t, x) \leqslant M$ (Proof of the second assertion leads to the first when we make a change in the sign of u.)

Suppose that ε is any positive number. Let us show that $u(t_0, x_0) \leqslant M + \varepsilon$ at an arbitrary point (t_0, x_0) of the strip S. Consider the function $w(t, x) = 2t + x^2$, which is a solution of equation (38.1). Let us set $N = \sup\limits_S |u(t, x)|$. The function

$$\frac{\varepsilon w(t, x)}{w(t_0, x_0)} + M - u(t, x),$$

which satisfies equation (38.1) everywhere in S is nonnegative for $t = 0$ and for

$$|x| = \sqrt{\frac{(N - M)\, w(t_0, x_0)}{\varepsilon}} + |x_0|,$$

since, for this value of $|x|$

$$\frac{\varepsilon w(t, x)}{w(t_0, x_0)} \geqslant \frac{\varepsilon x^2}{w(t_0, x_0)} \geqslant N - M.$$

From the extreme-value theorem for a bounded region (see subsection 2 of Section 38), this function must be nonnegative everywhere inside the rectangle

$$\left\{0 \leqslant t < T, |x| \leqslant \sqrt{\frac{(N - M)\, w(t_0, x_0)}{\varepsilon}} + |x_0|\right\},$$

in which the point (t_0, x_0) lies. Consequently, in this rectangle

$$u(t,x) \leqslant M + \frac{\varepsilon w(t, x)}{w(t_0, x_0)},$$

so that

$$u(t_0, x_0) \leqslant M + \varepsilon.$$

Since the point (t_0, x_0) and the number ε are arbitrary, it follows from the last inequality that $u(t, x) \leqslant M$ everywhere in the strip S.

Corollaries. (1) A bounded solution to the Cauchy problem for equation (38.1) in the strip S is unique.

(2) The solution to the Cauchy problem for equation (38.1) in the class of bounded functions depends continuosly on the initial condition given for $t=0$.

These assertions follow from the fact that the difference between two bounded solutions u_1 and u_2 of equation (38.1) in the strip S does not exceed in absolute value

$$\sup_{-\infty < x < \infty} |u_1(0, x) - u_2(0, x)|.$$

Remark: We have shown that the solution to the Cauchy problem in the class of bounded functions in unique. The following stronger assertion is also true:

Suppose that

$$f(x) = \max_{0 \leqslant t \leqslant T} |u(t, x)|.$$

If $u(t, x)$ satisfies equation (38.1) for $t > 0$, if $u(0, x) = 0$ for $-\infty < x < +\infty$ and if a constant C exists such that

$$f(x)\, e^{-Cx^2} \to 0 \text{ when } |x| \to \infty,$$

then $u(t, x) \equiv 0$.

This assertion is easily proven (we leave this to the reader) in the same way as we proved the uniqueness of the problem in question in the class of bounded functions if, instead of the function $w(t, x)$, we consider the function

$$W = 8(C+1)^2 t + e^{[8(C+1)^2 t + (C+1)]\, x^2},$$

which is positive for $t > 0$ and satisfies the condition $\frac{\partial^2 W}{\partial x^2} - \frac{\partial W}{\partial t} \leqslant 0$ for sufficiently small t. For functions satisfy-

ing the condition $\frac{\partial^2 W}{\partial x^2} - \frac{\partial W}{\partial t} \leqslant 0$, the theorem on minima is valid.

A.N. Tikhonov* has exhibited solutions to equation (38.1) for arbitrary $\varepsilon > 0$ that are not identically equal to zero but for which $u(0, x) = 0$ for $-\infty < x < +\infty$ and $f(x) e^{-x^{2+\varepsilon}} \to 0$ for $|x| \to \infty$.

3. Let us show that the solution to our problem is given by the formulae

$$u(t, x) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\varphi(\xi)}{\sqrt{t}} e^{-\frac{(x-\xi)^2}{4t}} d\xi \text{ when } t > 0, \tag{40.1}$$

$$u(0, x) = \varphi(x). \tag{40.2}$$

The integral in (40.1) is called Poisson's integral.

It is easy to verify that the integral (40.1) converges for all positive values of t. Similarly, it is easy to verify that the integrals obtained from (40.1) by differentiating under the integral sign with respect to t and x arbitrarily many times converge. Furthermore, all these integrals converge uniformly in a neighbourhood of an arbitrary point (t, x) if $t > 0$. From this it follows that for $t > 0$, there exists a function $u(t, x)$ defined by formula (40.1) that has derivatives with respect to t and to x of all orders. Since the integrand satisfies equation (38.1) for positive values of t, it follows that the function $u(t, x)$ itself satisfies this equation for positive values of t.

Let us show that the function $u(t, x)$ defined by formula (40.1) is bounded for $t > 0$. Note that if

$$M_1 = \max_{-\infty < x < \infty} |\varphi(x)|,$$

we have

$$|u(t, x)| \leqslant \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{M_1}{\sqrt{t}} e^{-\frac{(x-\xi)^2}{4t}} d\xi = \frac{M_1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\eta^2} d\eta = M_1.$$

It remains to show that the function $u(t, x)$ that we have defined is continuous for $t = 0$, that is, that, for every x_0,

* *Matem. sbornik*, **42**: 2, 199-216 (1935).

$$\left|\frac{1}{2\sqrt{\pi}}\int\limits_{-\infty}^{\infty}\frac{\varphi(\xi)}{\sqrt{t}}e^{-\frac{(x-\xi)^2}{4t}}d\xi-\varphi(x_0)\right|<\varepsilon, \tag{40.3}$$

if t and $|x-x_0|$ are sufficiently small. Note first of all that for this it will be sufficient for us to show that

$$\left|\frac{1}{2\sqrt{\pi t}}\int\limits_{-\infty}^{\infty}\varphi(\xi)\,e^{-\frac{(x-\xi)^2}{4t}}d\xi-\varphi(x)\right|<\frac{\varepsilon}{2} \tag{40.4}$$

for sufficiently small t and arbitrary x in some neighbourhood of the point x_0 since $|\varphi(x)-\varphi(x_0)|$ is small when $|x-x_0|$ is sufficiently small (because of the continuity of the function $\varphi(x)$).

To prove (40.4), we rewrite Poisson's integral (40.1), setting

$$\xi=x+2\sqrt{t}\,\zeta,$$

so that we have

$$\frac{1}{\sqrt{\pi}}\int\limits_{-\infty}^{\infty}\varphi(x+2\sqrt{t}\,\zeta)\,e^{-\zeta^2}d\zeta$$

and we note that

$$\varphi(x)=\frac{1}{\sqrt{\pi}}\int\limits_{-\infty}^{\infty}\varphi(x)\,e^{-\zeta^2}d\zeta.$$

Since $\varphi(x)$ is bounded, the integrals

$$\int\limits_{-\infty}^{-N}\varphi(x+2\sqrt{\;}\zeta)\,e^{-\zeta^2}d\zeta;\quad \int\limits_{-\infty}^{-N}\varphi(x)\,e^{-\zeta^2}d\zeta;$$

$$\int\limits_{N}^{\infty}\varphi(x+2\sqrt{t}\,\zeta)\,e^{-\zeta^2}d\zeta;\quad \int\limits_{N}^{\infty}\varphi(x)\,e^{-\zeta^2}d\zeta$$

can be made arbitrarily small in absolute value for sufficiently large N and arbitrary x. Therefore, for sufficiently large N and arbitrary x the following approximations are arbitrarily accurate:

$$u(t,x)\approx\frac{1}{\sqrt{\pi}}\int\limits_{-N}^{N}\varphi(x+2\sqrt{t}\,\zeta)\,e^{-\zeta^2}d\zeta,$$

and

$$\varphi(x) \approx \frac{1}{\sqrt{\pi}} \int_{-N}^{N} \varphi(x) e^{-\zeta^2} d\zeta.$$

But for sufficiently small t and arbitrary x in some neighbourhood of the point x_0, the right members of these approximations are arbitrarily close to each other because of the continuity of $\varphi(x)$. Inequality (40.4) then follows.

4. Thus, we have shown that the solution given by equations (40.1) and (40.2) is the only bounded solution of the problem posed at the beginning of the present section.

In particular, it follows from these formulae that if $\varphi(x)$ is equal to zero everywhere except for an arbitrarily small interval of values of x on which it is positive, the solution $u(t, x)$ will be positive for all values of x and any fixed positive value of t. From this follows the paradoxical assertion that heat flows in a rod with infinite speed. Of course, this is physically impossible. However, such a conclusion follows inexorably if we assume that the flow of heat in a rod is exactly described by equation (38.1). Obviously, the hypotheses taken in our derivation of this equation are not borne out exactly in experiment.

However, experiment shows that equation (38.1) does give a sufficiently good approximate description of the actual physical process of heat flow.

Problem 1. Suppose that $u(t, x)$ is a bounded solution of the Cauchy problem for the heat-flow equation (38.1) in the half-plane $t > 0$. Show that, for arbitrary x,

$$\lim_{t \to \infty} u(t, x) = \frac{a+b}{2},$$

$$\lim_{x \to -\infty} u(0, x) = a \text{ and } \lim_{x \to +\infty} u(0, x) = b.$$

2. Show that the Cauchy problem for the heat-flow equation (38.1) with an initial condition given for $t = 0$ is correctly stated in an arbitrary strip

$$\{-T < t < 0, \ -\infty < x < \infty\}.$$

41. SUMMARY OF SOME MORE EXTENSIVE INVESTIGATIONS OF PARABOLIC EQUATIONS

1. The existence and uniqueness of solutions to the first boundary-value problem for the equation

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x_1^2}+\ldots+\frac{\partial^2 u}{\partial x_n^2} \tag{41.1}$$

have been proven for arbitrary n. This problem is formulated in the simplest case as follows:

We are seeking a continuous function $u(t, x_1, \ldots, x_n)$ that is defined on a closed region $\bar{G}$ bounded "above" and "below" by portions of the hyperplanes $t=0$ and $t=T$ and on the sides by one or several surfaces with a continuously turning tangent hyperplane that nowhere intersects the t-axis. The function u must satisfy equation (41.1) within $\bar{G}$ and must coincide with some function f given on the lateral surface S and on the lower base $t=0$. This function is assumed to be continuous on the entire closed set on which it is defined. The uniqueness of the solution of the problem and its continuous dependence on the function f are proven in the same way as this was done in Section 38.

The first boundary-value problem can also be posed for regions of a more general form. Conditions have been found that must be satisfied by the boundary of the region in order for the first boundary-value problem to have a solution*.

2. The linear second-order parabolic equation with two independent variables

$$\frac{\partial u}{\partial t}=A(t, x)\frac{\partial^2 u}{\partial x^2}+B(t, x)\frac{\partial u}{\partial x}+C(t, x)u+D(t, x), \tag{41.2}$$

where A, B, C, and D are bounded functions and $A(t, x) \geqslant \alpha > 0$ can, by the substitution $u=ve^{Kt}$, be transformed to the form

$$\frac{\partial v}{\partial t}=A(t, x)\frac{\partial^2 v}{\partial x^2}+B(t, x)\frac{\partial v}{\partial x}+C_1(t, x)v+D_1(t, x). \tag{41.3}$$

Here, $C_1=C-K$ and $D_1=De^{-Kt}$. Let us assume that

* PETROVSKII, I.G., *Compositio Mathematica,* 1:3, 383-419 (1935).

$$K > M_1 \equiv \sup |C(t, x)|;$$

Then,

$$C_1(t, x) \leqslant M_1 - K < 0.$$

Suppose that the function $v(t, x)$ satisfies equation (41.3) within a curvilinear quadrilateral G bounded by the straight line segments $t=0$ and $t=T$ and by the curves $x=\varphi_1(t)$ and $x=\varphi_2(t)$ (where $\varphi_1(t) < \varphi_2(t)$), and that it coincides with some continuous function f on the lower base and on the lateral sides of G. Then, everywhere in G, we have

$$|v(t, x)| \leqslant \max\left\{M, \frac{M_2}{K-M_1}\right\}, \tag{41.4}$$

where $M = \max|f|$ and $M_2 = \sup|D(t, x)|$.

This is true because, if $v(t, x)$ assumes its greatest positive value at some point lying either on the upper surface of or within G, then $\frac{\partial v}{\partial t} \geqslant 0$, $\frac{\partial v}{\partial x} = 0$, $\frac{\partial^2 v}{\partial x^2} \leqslant 0$, and $C_1 v < 0$ at that point. From equation (41.3) we know that in this case $\max v(t, x) \leqslant \frac{M_2}{K-M_1}$. On the other hand, if v assumes its largest value on the lower base or on the lateral surface of G, we shall have $v \leqslant M$ everywhere in G. Thus

$$v \leqslant \max\left\{M, \frac{M_2}{K-M_1}\right\}.$$

Analogously, we can show that

$$v \geqslant -\max\left\{M, \frac{M_2}{K-M_1}\right\}.$$

For the function $u = ve^{Kt}$ satisfying equation (41.3) within G, we obtain from (41.4) the inequality

$$|u(t, x)| \leqslant \max\left\{Me^{KT}, \frac{M_2 e^{KT}}{K-M_1}\right\}$$

which is a generalisation of the extreme-value theorem (see Section 38).

An analogous assertion holds for solutions $u(t, x_1, \ldots, x_n)$ of the parabolic equation

$$\frac{\partial u}{\partial t} = \sum_{i, j=1}^{n} A_{ij}(t, x_1, \ldots, x_n) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} B_i(t, x_1, \ldots, x_n) \frac{\partial u}{\partial x_i}$$
$$+ C(t, x_1, \ldots, x_n) u + D(t, x_1, \ldots, x_n), \quad (41.5)$$

where the form

$$\sum_{i, j=1}^{n} A_{ij}(t, x_1, \ldots, x_n) \alpha_i \alpha_j$$

is positive definite at all points of the region in question.

3. The first boundary-value problem for the parabolic equation (41.5) in a region G bounded by portions of the hyperplanes $t=0$ and $t=T$ and by the surface S has a unique solution for an arbitrary continuous function f given on S for $t=0$ if (1) the coefficients A_{ij}, B_i, C and D satisfy a Holder* condition in the region G and (2) every point P of the surface S touches a sphere with centre Q at points of which except P lie outside $\bar{G}$ and the straight line QP is not parallel to the t-axes**.

4. Consider the quasilinear second-order parabolic equation with two independent variables

$$\frac{\partial u}{\partial t} = A(t, x, u) \frac{\partial^2 u}{\partial x^2} + B(t, x, u) \frac{\partial u}{\partial x} + C(t, x, u),$$

where A, B, and C have continuous derivatives of sufficiently high orders, where $A(t, x, u) \geqslant \alpha > 0$ and $C'_u(t, x, u) < c$ (here, α and c are constants). The existence and uniqueness of a solution to the first boundary-value problem in an arbitrary rectangle $\{0 \leqslant t \leqslant T,\ 0 \leqslant x \leqslant l\}$, and also the existence and uniqueness of a bounded solution to the Cauchy problem in an arbitrary strip $\{0 \leqslant t \leqslant T,\ -\infty < x < \infty\}$ have been proven for this equation. These problems have also been studied for quasilinear parabolic equations with a

* See footnote Section 37, subsection 2.

** FRIEDMAN, *Journal of Mathematics and Mechanics*, 7, No. 5, 771-791 (1958).

greater number of independent variables*.

5. The system of linear equations

$$\frac{\partial u_i}{\partial t} = \sum_{j=1}^{N} \sum_{0 \leqslant k_1 + \ldots + k_n \leqslant 2m} A_{ij}^{(k_1 \ldots k_n)}(t, x_1, \ldots, x_n) \frac{\partial^k u_j}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} + F_i(t, x_1, \ldots, x_n) \quad (i = 1, \ldots, N) \tag{41.6}$$

is said to be parabolic at the point $(t^{(0)}, x_1^{(0)}, \ldots, x_n^{(0)})$ if for arbitrary real numbers $\alpha_1, \ldots, \alpha_n$ the sum of the squares of which is equal to 1, all the roots $\lambda_1, \ldots, \lambda_N$ of the determinant

$$\left| \sum_{k_1 + \ldots + k_n = 2m} (-1)^m A_{ij}^{(k_1 \ldots k_n)} \, t^{(0)}, \ldots, x_n^{(0)}) \, \alpha_1^{k_1} \ldots \alpha_n^{k_n} - \lambda \delta_{ij} \right|$$

have negative real parts.

For parabolic systems, the Cauchy problem is correctly stated for positive t in the class of bounded functions with sufficiently smooth initial conditions for $t = 0$. Correctness of the statement of the problem is maintained in the class of functions that increase as $x_1^2 + \ldots + x_n^2 \to \infty$ no faster than

$$e^{\left(x_1^2 + \ldots + x_n^2 \right)^{\frac{m}{2m-1}}},$$

where 2 is the order of the system.

If all the coefficients $A_{ij}^{(k_1 \ldots k_n)}$ and F_i of the parabolic system (41.6) are analytic with respect to the arguments $x_1, \ldots, x_n$ all sufficiently smooth solutions of this system are also analytic with respect to $x_1, \ldots, x_n$ †

* See *Mathematics in the USSR in the last forty years*, **I**, FIZMATGIZ 604-628 (1959).

† See the preceding footnote.

Supplement

42. SOLUTION OF THE FIRST BOUNDARY-VALUE PROBLEM FOR HEAT-FLOW EQUATION USING NETS

1. We shall prove the existence of a solution to the first boundary-value problem for equation (38.1) by the method of nets. This proof will also indicate a method of constructing an approximate solution of the problem.

Let us denote by G the region bounded by the straight line segments $t=t_0$ and $t=T$ (where $T>t_0$) and the curves $x=\varphi_1(t)$ and $x=\varphi_2(t)$. It is assumed that $\varphi_1(t)$ and $\varphi_2(t)$ are continuous and that $\varphi_1(t)<\varphi_2(t)$ for $t_0 \leqslant t \leqslant T$. We denote by Γ the lower base $(t=t_0)$ and the lateral sides $(x=\varphi_1(t)$, $x=\varphi_2(t)$, and $t_0 \leqslant t \leqslant T)$ of the curvilinear quadrilateral G. We are seeking a solution to equation (38.1) in G that will be continuous in G and on its boundary and that will assume on Γ the values of a continuous function f. For the first boundary-value problem to be solvable, the function $\varphi_1(t)$ and $\varphi_2(t)$ must satisfy certain supplementary conditions, which we shall state below.

On the tx-plane containing the region G, let us draw two families (a net) of lines parallel to the coordinate axes:

$$x=mh, \quad t=t_0+kh,$$

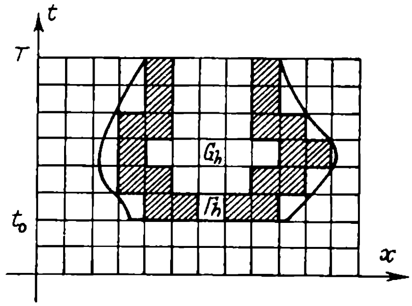
where h is some positive number and m and k assume successive integral values, so that the entire region G is

covered by squares of side h. We shall call the corners of these squares nodes or nodal points of the net that we have constructed.

Let us denote by $\bar{G}_h$ the set of those squares that belong entirely to $G \cup \Gamma$. We denote by Γ_h the set of all these squares belonging to $\bar{G}_h$ at least one corner of which belongs to the boundary of $\bar{G}_h$ except the interior squares of the top layer (that is, except the squares at least one corner of which have the maximum t-coordinate of points in $\bar{G}_h$ and whose lateral edges do not belong to the boundary of $\bar{G}_h$. We denote by G_h the nodal points belonging to $\bar{G}_h - \Gamma_h$ (see Fig. 19).

In each node of Γ_h we define a function f_h that is equal to the value of the function f at the point of Γ closest to

Fig. 19



that node or at one of these points if there are several of them. Let us compare equation (38.1) with the following difference equation*:

$$\frac{u(t, x) - u(t-h, x)}{h} = \frac{u(t, x+h) - 2u(t, x) + u(t, x-h)}{h^2}. \quad (42.1)$$

We shall seek a function u_h defined at the nodal points $\bar{G}_h$ that satisfies equation (42.1) at the nodes (t, x) in G_h and that coincides with f_h at the nodal points of the Γ_h. Let us show that a unique function u_h exists satisfying these conditions.

2. Lemma. A function u_h defined at the nodes $\bar{G}_h$ and satisfying equation (42.1) at the nodal points (t, x) in G_h has maximum and minimum values at the nodal points of Γ_h.

Proof: Let us assume that the function u_h assumes at

* If the function $u(t, x)$ has derivatives $\partial u/\partial t$ and $\partial^2 u/\partial x^2$ at the point (t, x), equation (42.1) becomes equation (38.1) when we pass to the limit as $h \to 0$.

certain nodes of G_h values of u_h at the nodal points of Γ_h. In this case, there is a nodal point (t_1, x_1) in G_h at which u_h assumes its greatest value; also, at least one of the adjacent points, the value of u_h is less than $u_h(t_1, x_1)$. By adjacent points to (t_1, x_1), we mean the nodal points

$$(t_1 - h, x_1),\ (t_1, x_1 + h), \quad (t_1, x_1 - h).$$

If $u_h(t_1 - h, x_1) < u_h(t_1, x_1)$, the left side of equation (42.1) will be positive at the point (t_1, x_1) and the right side will be nonpositive. This leads us to a contradiction since equation (42.1) must be satisfied at the point (t_1, x_1). Similarly, we shall arrive at a contradiction if we assume that

$$u_h(t_1, x_1 + h) < u_h(t_1, x_1)$$

or

$$u_h(t_1, x_1 - h) < u_h(t_1, x_1).$$

In an analogous manner, we can prove that u_h cannot assume values in G_h that are smaller than the smallest value of u_h on Γ_h.

By using this lemma, let us show that for an arbitrary function f_h given at the nodes of Γ_h, there exists a unique function u_h satisfying equation (42.1) at the nodal points G_h and coinciding with f_h on Γ_h. The values of u_h at the nodal points of G_h satisfy the linear algebraic system of equations that we obtain by writing equation (42.1) for each nodal point (t, x) in G_h. The number of equations of such a system will be equal to the number of unknowns u_h. The determinant of this system will be nonzero since the corresponding homogeneous system obtained by setting $f_h = 0$ at all the nodal points Γ_h has only a trivial solution (because of the above lemma). Consequently, the function u_h is uniquely defined.

3. We introduce the following notations:

$$u_{\bar{t}} = \frac{u(t, x) - u(t - h, x)}{h}, \quad u_x = \frac{u(t, x + h) - u(t, x)}{h},$$

$$u_{\bar{x}} = \frac{u(t, x) - u(t, x - h)}{h}.$$

In these notations, equation (42.1) takes the form

$$u_{\bar{t}} = u_{x\bar{x}}. \tag{42.2}$$

Suppose that $h_n = \frac{1}{2^n}$ and $u_{h_n} = u^n$ (for $n = 1, 2, \ldots$). We shall prove the existence of a solution to the first boundary-value problem in the following manner. First, we shall show that each function in the families $\{u^n\}$, $\{u^n_{\bar{t}}\}$, $\{u^n_{x\bar{x}}\}$ given on the nets can be extended to all points of the region G in such a way that in an arbitrary subregion G^* whose closure is contained in G, the families of functions $\{u^n\}$, $\{u^n_{\bar{t}}\}$ and $\{u^n_{x\bar{x}}\}$ will be uniformly bounded and equicontinuous. Then, by applying Arzelà's theorem†, we shall show that the sequence $\{u^n\}$ contains a subsequence $\{u^{\bar{n}}\}$ that converges uniformly G to some function $u(t, x)$ in each region G^* whose closure is contained in G. Here, the sequence $\{u^{\bar{n}}_{\bar{t}}\}$ converges to $\frac{\partial u}{\partial t}$, and $\{u^{\bar{n}}_{x\bar{x}}\}$ converges to $\frac{\partial^2 u}{\partial x^2}$. Taking the limit as $h_n \longrightarrow 0$ in equation (42.2), we see that the limiting function $u(t, x)$ satisfies equation (38.1) in G.

Finally, with the aid of barrier functions, we can show just as was done in Section 31 that the limiting function $u(t, x)$ is continuous in $G \cup \Gamma$ and that it assumes on Γ the values of the given continuous function f. By using the uniqueness of the solution of the first boundary-value problem, we can show that not only the subsequence $\{u^{\bar{n}}\}$ but also every other sequence $\{u^n\}$ converges to $u(t, x)$ and consequently, that the solutions u_{h_n} of equation (42.1) represent the solution of the corresponding first boundary-value problem with an arbitrary degree of accuracy provided h_n is sufficiently small.

4. It follows from the lemma of subsection 2 that $|u^n| \leqslant \max|f|$ for arbitrary n. Let us show that uniform boundedness of the family $\{u^n\}$ in G implies uniform boundedness of the family $\{u^n_x\}$ in every region G^* whose closure is contained in the boundary of G.

Here, it will be sufficient for us to show that this assertion is valid for a rectangle Q with sides parallel to the coordinate axes since an arbitrary region G^* can be covered by a finite number of such rectangles belonging to G.

To get an estimate for u^n_x in Q, we use a device employed

† PETROVSKII, I.G., *Lectures on the theory of ordinary differential equations*, Gostekhizdat, 40 (1952).

by Bernstein for estimating the derivatives of the solution of a parabolic equation†.

Without loss of generality, we may assume that the rectangle Q is defined by the inequalities

$$|x| \leqslant a, \quad 0 \leqslant t \leqslant b$$

and that its sides belong to a net beginning with sufficiently large n. In the following calculations of this subsection, we shall omit, for simplicity of writing, the index n of the function u^n.

Let us consider the function

$$z = u_x^2 F + Cv, \tag{42.3}$$

at the nodal points of the rectangle Q, where

$$F = t(a^2 - x^2)^2, \; v = u^2(t, x+h) + u^2(t, x-h) + u^2(t-h, x),$$

and C is some positive constant. Let us show that if C is sufficiently great, $z(t, x)$ will assume its greatest value either at $t = 0$ or on the sides $x = \pm a$ of the rectangle Q. It will then be easy to obtain an estimate for $u_x(t, x)$. For if $|u(t, x)| \leqslant M$ in G, then $z \leqslant 3CM^2$ on the sides $t = 0$ and $x = \pm a$ on the rectangle Q and, consequently, $z \leqslant 3CM^2$ everywhere in Q. So that

$$u_x^2 \leqslant \frac{3CM^2}{t(a^2 - x^2)^2},$$

that is, the u_x are uniformly bounded in a rectangle Q^* lying within Q.

To show that $z(t, x)$ attains its greatest value at $t = 0$ or $x = \pm a$, for sufficiently large C, let us calculate the quantity

$$L(z) \equiv z_{\overline{x}x} - z_{\overline{t}}.$$

We shall show that, for sufficiently large C,

$$L(z) \geqslant 0 \text{ in } Q.$$

By use of the same reasoning as we used in proving the lemma of subsection 2, we see from this inequality that $z(t, x)$ has its largest value on the boundary $t = 0$ or $x = \pm a$ of the rectangle Q.

To calculate $L(z)$, we use the formula

† BERNSTEIN, S.N., *Dokl. Akad. Nauk SSSR*, **18**, No. 7, 385-388 (1938).

$$L(\varphi\psi) = \varphi L(\psi) + \psi L(\varphi) + h\varphi_{\bar t}\psi_{\bar t} + \varphi_x\psi_x + \varphi_{\bar x}\psi_{\bar x}, \quad (42.4)$$

the validity of which is easily established for arbitrary functions φ and ψ given on the nets. We have

$$L(z) = u_x^2 L(F) + FL(u_x^2) + hF_{\bar t}(u_x^2)_{\bar t} + F_x(u_x^2)_x + F_{\bar x}(u_x^2)_{\bar x} + CL(v).$$

By using formula (42.4), we obtain

$$L(u_x^2) = 2u_x L(u_x) + h(u_{x\bar t})^2 + u_{xx}^2 + u_{x\bar x}^2 = u_{xx}^2 + u_{x\bar x}^2 + hu_{x\bar t}^2,$$

since $L(u_x) = 0$.

It is easy to see that

$$\begin{aligned} L(u^2(t, x+h)) &= 2u(t, x+h)\, L(u(t, x+h)) + u_x^2(t, x+h) \\ &\quad + u_{\bar x}^2(t, x+h) + hu_{\bar t}^2(t, x+h) \\ &= u_x^2(t, x+h) + u_{\bar x}^2(t, x+h) + hu_{\bar t}^2(t, x+h) \\ &= u_x^2(t, x+h) + u_x^2(t, x) + hu_{\bar t}^2(t, x+h), \end{aligned}$$

$$\begin{aligned} (u_x^2)_x &= [u_x + u_x(t, x+h)]\, u_{xx}, \\ (u_x^2)_{\bar x} &= [u_x + u_x(t, x-h)]\, u_{x\bar x}, \\ h(u_x^2)_{\bar t} &= u_x^2 - u_x^2(t-h, x). \end{aligned}$$

Consequently,

$$\begin{aligned} L(z) &= C[u_x^2(t, x+h) + u_x^2(t, x) + hu_{\bar t}^2(t, x+h) + u_x^2(t, x-h) \\ &+ u_{\bar x}^2(t, x-h) + hu_{\bar t}^2(t, x-h) + u_x^2(t-h, x) + u_{\bar x}^2(t-h, x) \\ &+ hu_{\bar t}^2(t-h, x)] + u_x^2 L(F) + F_{\bar t}[u_x^2 - u_x^2(t-h, x)] \\ &+ F(u_{xx}^2 + u_{x\bar x}^2 + hu_{x\bar t}^2) + F_x[u_x + u_x(t, x+h)]\, u_{xx} \\ &+ F_{\bar x}[u_x + u_x(t, x-h)]\, u_{x\bar x}. \end{aligned}$$

Let us estimate the terms

$$Fu_{xx}^2 + F_x[u_x + u_x(t, x+h)]\, u_{xx}.$$

Since $\quad F_x = -2t(2x+h)(a^2 - x^2) + ht(2x+h)^2,$

we have $\quad Fu_{xx}^2 + F_x[u_x + u_x(t, x+h)]\, u_{xx}$

$$\begin{aligned} &= t\{(a^2 - x^2)\, u_{xx} - (2x+h)[u_x + u_x(t, x+h)]\}^2 \\ &\quad - t(2x+h)^2[u_x + u_x(t, x+h)]^2 \\ &\quad + ht(2x+h)^2[u_x + u_x(t, x+h)]\, u_{xx} \end{aligned}$$

$$\geqslant -t(2x+h)^2[u_x+u_x(t,x+h)]^2$$
$$+t(2x+h)^2[u_x^2(t,x+h)-u_x^2] =$$
$$= -t(2x+h)^2[2u_x^2+2u_x\cdot u_x(t,x+h)]$$
$$\geqslant -t(2x+h)^2[3u_x^2+u_x^2(t,x+h)].$$

Analogously, we obtain

$$Fu^2_{x\bar{x}}+F_{\bar{x}}\,[u_x+u_x(t,x-h)]\,u_{x\bar{x}}$$
$$\geqslant -t(2x-h)^2[3u_x^2+u_x^2(t,x-h)].$$

By using the last inequalities and the facts that $F>0$ and $F_{\bar{t}}>0$,we obtain

$$L(z)\geqslant C\,[u_x^2(t,x+h)+u_x^2(t,x)+u_x^2(t,x-h)+u_x^2(t-h,x)]$$
$$+L(F)\,u_x^2-F_{\bar{t}}\,u_x^2(t-h,x)-t(2x+h)^2[3u_x^2+u_x^2(t,x+h)]$$
$$-t(2x-h)^2[3u_x^2+u_x^2(t,x-h)]$$
$$=[C-L(F)-3t(2x+h)^2-3t(2x-h)^2]\,u_x^2$$
$$+[C-t(2x+h)^2]\,u_x^2(t,x+h)+[C-t(2x-h)^2]\,u_x^2(t,x-h)$$
$$+[C-F_{\bar{t}}]\,u_x^2(t-h,x). \qquad (42.5)$$

Obviously, if C is sufficiently great, all the terms on the right side of inequality (42.5) will be nonnegative.(The constant C can be chosen independently of h.) Thus, we have shown that if the solutions $u^n(t,x)$ of equation (42.2) are uniformly bounded in G, the functions u_x^n will also be uniformly bounded in every region G^* whose closure lies within G. Since u_x^n, $u^n_{x\bar{x}}$, and $u^n_{x\bar{x}x}$, are also solutions to equation (42.2), we have proven the uniform boundedness of $u^n_{x\bar{x}}=u^n_{\bar{t}}$, $u^n_{x\bar{x}x}=u^n_{\bar{t}x}$ and $u^n_{x\bar{x}x\bar{x}}=u^n_{\bar{t}\bar{t}}$ in an arbitrary region G^* whose closure is contained in G.

5. Let us now show that the functions u^n can be extended to the entire region G in such a way that the family $\{u^n\}$ will be uniformly bounded and equicontinuous in any region G^* whose closure is contained in G.

To do this, we partition each square into triangles with a diagonal parallel to the straight line $t=x$. In each such triangle, we set u^n equal to a linear function that assumes at the corners of the triangle the values of u^n that were determined above. It is easy to see that the function u^n con-

structed in this way is continuous in $\bar{G}_{h_n}$, and that, within the triangle and on its sides, it cannot assume values greater or less than its values at the vertices of the triangle. At points of G that do not belong to $\bar{G}_{h_n}$, we define the function u^n arbitrarily except that we require that it be continuous and bounded in G.

It follows from the uniform boundedness of $u^n_{\bar{x}}$ and $u^n_{\bar{t}}$ in G^* that, for the nodal points (t, x) belonging to G^*,

$$|u^n(t, x + h_n) - u^n(t, x)| \leq Kh_n$$

and

$$|u^n(t, x) - u^n(t - h_n, x)| \leq Kh_n,$$

where K does not depend on n.

It will be sufficient to prove the equicontinuity of the functions u^n within the rectangle Q belonging to G with sides parallel to the coordinate axes since an arbitrary region G^* whose closure is contained in G can be covered by a finite number of rectangles of this form.

For any two nodal points (t_1, x_1) and (t_2, x_2) in Q, we have

$$|u^n(t_1, x_1) - u^n(t_2, x_2)| \leq 2K\sqrt{(t_1 - t_2)^2 + (x_1 - x_2)^2}.$$

Suppose that (t_1, x_1) and (t_2, x_2) are arbitrary points in Q and that (t_1^*, x_1^*) and (t_2^*, x_2^*) are the nodal points closest to (t_1, x_1) and (t_2, x_2) respectively. Because of the way in which the function u^n is defined, we have

$$|u^n(t_i^*, x_i^*) - u^n(t_i, x_i)| \leq 2Kh_n \, (i = 1, 2).$$

for sufficiently large n. Therefore, for sufficiently large n,

$$\begin{aligned} &|u^n(t_1, x_1) - u^n(t_1, x_2)| \\ &\quad \leq |u^n(t_1^*, x_1^*) - u^n(t_1, x_1)| + |u^n(t_2^*, x_2^*) - u^n(t_2, x_2)| \\ &\qquad\qquad + |u^n(t_1^*, x_1^*) - u^n(t_2^*, x_2^*)| \\ &\quad \leq 4Kh_n + 2K\sqrt{(t_1^* - t_2^*)^2 + (x_1^* - x_2^*)^2}. \end{aligned} \tag{42.6}$$

This inequality and the uniform continuity of each of the functions u^n in G^* imply the equicontinuity of the functions u^n in Q and, consequently, in G^*.

By applying Arzela's theorem, we obtain the result that the sequence of functions u^n contains a subsequence that converges uniformly in G^*.

In just the same way, we can show by using the uniform

boundedness (proven above) of $u^n_{\bar{t}}$, $u^n_{\bar{t}x}$, $u^n_{\bar{t}\bar{t}}$ u^n_x, $u^n_{\overline{xx}}$, and $u^n_{x\bar{t}}$, that each of the functions $u^n_{\bar{t}}$ and u^n_x can be extended to the entire region G in such a way that the families of functions $\{u^n_{\bar{t}}\}$ and $\{u^n_x\}$ will be uniformly bounded and equicontinuous in

Therefore, by Arzelà's theorem, an arbitrary infinite set of functions u^n contains a subsequence $\{u^{n'}\}$ that converges uniformly in G^* such that the sequences $\{u^{n'}_{\bar{t}}\}$ and $\{u^{n'}_x\}$ corresponding to it also converge uniformly in G^*.

Consider a sequence of regions G^*_m such that (1)

$$G^*_m \subset G^*_{m+1},$$

(2)

$$\sum_{m=1}^{\infty} G^*_m = G$$

and, (3) the closure of G^*_m is contained in G. Let us choose from $\{u^n\}$ a sequence $u^{11}, u^{12}, \ldots, u^{1k}, \ldots$ that converges uniformly in G^*_1 such that the sequences $\{u^{1k}_{\bar{t}}\}$ and $\{u^{1k}_x\}$ also converge uniformly in G^*_1. From the sequence $\{u^{1k}\}$ let us choose a subsequence

$$u^{21},\ u^{22}, \ldots, u^{2k}, \ldots \tag{42.7}$$

that converges uniformly in G^*_2 such that the subsequences $\{u^{2k}_{\bar{t}}\}$ and $\{u^{2k}_x\}$ converge in G^*_2, etc.

Consider the sequence of functions

$$u^{11},\ u^{22}, \ldots, u^{kk}, \ldots \tag{42.8}$$

It is easy to see that this sequence and the sequences $\{u^{kk}_{\bar{t}}\}$ and $\{u^{kk}_x\}$ converge at every point of the region G and that they converge uniformly in every region G^* whose closure is contained in G. We denote by $\{u^{kk}\}$, $\{u^{kk}_{\bar{t}}\}$ and $\{u^{kk}_x\}$ the limits (in G) of the sequences $U(t, x)$, $\bar{U}(t, x)$ and $\bar{\bar{U}}(t, x)$ respectively.

Let us show that

$$\frac{\partial U}{\partial t} = \bar{U}, \quad \frac{\partial U}{\partial x} = \bar{\bar{U}} \text{ and } \frac{\partial^2 U}{\partial x^2} = \bar{U}. \tag{42.9}$$

Suppose that the points (t_1, x) and (t_2, x) are nodal points of the net for sufficiently small h_n, that $t_1 - t_2 = l_k h_{kk}$ and that the line segment connecting them is contained in G.

Then,

$$u^{kk}(t_1, x) - u^{kk}(t_2, x) = \sum_{i=0}^{l_k - 1} u_{\bar{t}}^{kk}(t_1 - ih_{kk}, x)\, h_{kk}$$
$$= \sum_{i=0}^{l_k - 1} \overline{U}(t_1 - ih_{kk}, x)\, h_{kk} + \varepsilon_k, \quad (42.10)$$

where ε_k approaches zero as $k \longrightarrow \infty$ since the sequence $u_{\bar{t}}^{kk}$ converges uniformly to $\overline{U}$. If we pass to the limit in equation (42.10) as $k \longrightarrow \infty$, we obtain

$$U(t_1, x) - U(t_2, x) = \int_{t_2}^{t_1} \overline{U}\, dt. \quad (42.11)$$

Since the nodal points form an everywhere dense set in G and since the functions $U(t, x)$ and $\overline{U}(t, x)$ are continuous in G, equation (42.11) is valid for arbitrary points (t_1, x) and (t_2, x) if the segment connecting them is contained in G. Therefore, $\frac{\partial U}{\partial t} = \overline{U}$ everywhere in G. In just the same way we can show that

$$\frac{\partial U}{\partial x} = \overline{\overline{U}} \text{ and } \frac{\partial U(t, x_1)}{\partial x} - \frac{\partial U(t, x_2)}{\partial x} = \int_{x_2}^{x_1} \overline{U}\, dx,$$

if the points (t, x_1) and (t, x_2) belong to G, that is, if $\frac{\partial^2 U}{\partial x^2} = \overline{U}$ in G. Thus, we have shown that the limiting function $U(t, x)$ possesses first and second partial derivatives with respect $\frac{\partial U}{\partial t}$ and $\frac{\partial^2 U}{\partial x^2}$, and also $\frac{\partial U}{\partial t} = \frac{\partial^2 U}{\partial x^2}$ at all points of the region G.

6. Let us now study the behaviour of the limiting values of the function $U(t, x)$ on the boundary Γ of the region G.

Lemma 1. Suppose that the point A with coordinates (t_0, x_0) lies on the lower base $(t = t_0)$ of the curvilinear quadrilateral G. Then,

$$\lim_{(t,x) \to A,\ (t,x) \subset G} U(t, x) = f(A).$$

Proof: Consider the function

$$w = (x - x_0)^2 + 3(t - t_0).$$

At all points of $G \cup \Gamma$ other than A, the function $w(t, x)$ is

positive. It is easy to verify that

$$L(w) = w_{x\bar{x}} - w_{\bar{t}} < 0.$$

Let ε denote any arbitrarily small positive number. We denote by Ω_ε a sufficiently small neighbourhood of the point A that $|f_h - f(A)| < \varepsilon$ for values of f_h at all the nodal points of Γ_h that belong to Ω_ϵ for sufficiently small h. Let C denote a constant such that $Cw > 2 \max_\Gamma |f|$ at all points of $G \cup \Gamma$ not belonging to Ω_ϵ. Let us consider the function u^n only at the nodal points of $\overline{G}_{h_n}$. It is easy to show that at the nodes of $\overline{G}_{h_n}$

$$f(A) - \varepsilon - Cw(t, x) \leqslant u^n(t, x) \leqslant f(A) + \varepsilon + Cw(t, x). \quad (42.12)$$

Specifically, the functions

$$\varphi = f(A) - \varepsilon - Cw - u^n \text{ and } \psi = -f(A) - \varepsilon - Cw + u^n$$

are nonpositive at all nodal points of Γ_{h_n} because of the definition of Ω_ϵ and the choice of the constant C. Since $L(\varphi) > 0$ and $L(\psi) > 0$, the functions φ and ψ assume their greatest values on Γ_{h_n}. Consequently, at all nodal points of $\overline{G}_{h_n}$ the functions φ and ψ are nonpositive and inequalities (42.12) hold at all nodes belonging to $\overline{G}_{h_n}$.

If the points (t, x) is a nodal point of G_{h_n} for all sufficiently high n, then, by passing to the limit at that point in inequalities (42.12), we obtain

$$f(A) - \varepsilon - Cw \leqslant U(t, x) \leqslant f(A) + \varepsilon + Cw. \quad (42.13)$$

Since the set of points that are nodal of G_{h_n} for some sufficiently great n is everywhere dense in G and the function $U(t, x)$ is continuous in G, inequalities (42.13) are valid at all points of G. Consequently,

$$f(A) - \varepsilon \leqslant \varliminf_{(t, x) \to A} U(t, x) \leqslant \varlimsup_{(t, x) \to A} U(t, x) \leqslant f(A) + \varepsilon.$$

Since ε is an arbitrary positive number,

$$\lim_{(t, x) \to A} U(t, x) = f(A),$$

which completes the proof.

Lemma 2. Suppose that a point A with coordinates (t_1, x_1)

lies on a lateral side of the curvilinear quadrilateral G. Then,

$$\lim_{(t,\,x)\to A,\ (t,\,x)\subset G} U(t, x)=f(A)$$

if a (barrier) function $v_A(t, x)$ exists possessing the following properties: (1) $v_A(t, x)$ is defined and continuous at those points of the intersection of $G \cup \Gamma$ and some neighbourhood of A at which $t \leqslant t_1$. We denote by Ω_A the set of points at which v_A is defined. (2) $v_A(A)=0$ and $v_A(t, x)>0$ at all points of Ω_A other than A. (3) $L(v)\leqslant 0$ at all nodal points G_{h_n} that belong to Ω_A for sufficiently large n.

Proof: Suppose that the point A with coordinates (t_1, x_1) lies on the curve $x=\varphi_1(t)$. Let us choose $\alpha>0$ sufficiently small that the region D_α bounded by the straight lines $t=t_1$, $t=t_1-\alpha$, $x=x_1+\alpha$ and the curve $x=\varphi_1(t)$ will be contained in Ω_A.

Let ε denote an arbitrary positive number. We denote by Ω_ϵ a sufficiently small neighbourhood of the point A that

$$|f_h-f(A)|<\varepsilon \tag{42.14}$$

for values of f_h at all nodal points of Γ_h belonging to Ω_ε for sufficiently small h. Let C_1 denote a constant such that $C_1 v_A>2\max|f|$ at all points of D_α not belonging to Ω_ε. As in the preceding lemma, we see that at all the nodal points of $\overline{G}_h$ belonging to D_α, for sufficiently large n,

$$f(A)-\varepsilon-C_1 v_A\leqslant u^n\leqslant f(A)+\varepsilon+C_1 v_A. \tag{42.15}$$

Therefore, taking the limit as $n\to\infty$, and following the same reasoning as with Lemma 1, we obtain

$$\lim_{(t,\,x)\to A,\ t\leqslant t_1} U(t, x)=f(A). \tag{42.16}$$

It follows from inequalities (42.15) that there exists a neighbourhood Ω of the point A such that, at all nodal points of $\overline{G}_{h_n}$ belonging to Ω.

$$|u^n(t, x)-f(A)|<2\varepsilon, \tag{42.17}$$

if $t\leqslant t_1$. We may assume that Ω belongs to Ω_ε.

Consider the function $w=(x-x_1)^2+3(t-t_1)$. Let C_2 de-

note a constant such that $C_2 w > 2 \max_{\Gamma} |f|$ at all points of not belonging to Ω and located above the straight line $t = t_1 - \delta$, where δ is an arbitrary small positive number.

Just as with the preceding lemma it is easy to show that

$$f(A) - 3\varepsilon - C_2 w \leqslant u^n \leqslant f(A) + 3\varepsilon + C_2 w. \qquad (42.18)$$

at all nodal points of G_{h_n} lying above the straight line $t = t_1 - h_n$ for sufficiently small h_n. To show this, consider the nodal points at which $t > t_1 - h_n$. We denote the set of these points by H_n. We shall show that inequality (42.18) is satisfied at the nodal points of H_n belonging to Γ_n and at the nodes of H_n in the lowest layer. For nodes outside Ω, these inequalities are satisfied for sufficiently large n by virtue of the choice of C_2. For nodes belonging to Ω, these inequalities are satisfied because of inequalities (42.14) and (42.17) provided w is positive at these nodes. On the other hand, if w is negative at the nodal point in Ω in question, inequalities (42.18) are satisfied for sufficiently large because of the fact that $C_2 w > -\varepsilon$ for $t > t_1 - h_n$ if h_n is sufficiently small and because of inequality (42.17).

Since $L(w)$ is negative, we can show, as in the preceding lemma, that inequalities (42.18) are valid at all points of H_n without exception.

From these inequalities, by following the same reasoning as in the preceding lemma, we obtain

$$\lim_{(t,\, x) \to A,\ t \geqslant t_1} U(t, x) = f(A). \qquad (42.19)$$

The assertion of the lemma for points A lying on the curve $x = \varphi_1(t)$. follows from (42.16) and (42.19). For points A belonging to the curve $x = \varphi_2(t)$, the proof is analogous.

Theorem. The first boundary-value problem for the heat-flow equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

has a solution in the region G if its boundary Γ satisfies the following conditions: (1) For every point (t_1, x_1) lying on the curve $x = \varphi_1(t)$, there exists a positive number k_1 such that, for $t < t_1$ and sufficiently small $t_1 - t$

$$\varphi_1(t) - \varphi_1(t_1) > k_1(t - t_1).$$

(2) For each point (t_2, x_2) lying on the curve $x=\varphi_2(t)$, there exists a positive number k_2 such that, for $t<t_2$ and sufficiently small t_2-t

$$\varphi_2(t)-\varphi_2(t_2)<-k_2(t-t_2).$$

(In particular, conditions (1) and (2) will be satisfied if the functions $\varphi_1(t)$ and $\varphi_2(t)$ satisfy a Lipschitz condition.)

Proof: It will be sufficient to show that, under these conditions, the function that is equal to $f(A)$ on Γ and $U(t, x)$ in G is continuous in $G \cup \Gamma$.

Because of Lemmas 1 and 2, this function will be continuous in $G \cup \Gamma$ if, for every point A on the curves $x=\varphi_1(t)$ and $x=\varphi_2(t)$, there exists a (barrier) function v_A. If the point A with coordinates (t_1, x_1) lies on the curve $x=\varphi_1(t)$, the function

$$v_A(t, x)=\frac{1}{[(x_1-x')^2+(t_1-t')^2]^N}-\frac{1}{[(x-x')^2+(t-t')^2]^N}$$

may serve as a barrier, where

$$x'=x_1-\frac{\rho}{\sqrt{1+k_1^2}}, \quad t'=t_1+\frac{k_1\rho}{\sqrt{1+k_1^2}}, \quad \rho>0$$

ρ is a sufficiently small positive number, and N is a sufficiently large positive number. At the point (t_2, x_2) of the curve $x=\varphi_2(t)$, the function

$$v_A(t, x)=\frac{1}{[(x_2-x'')^2+(t_2-t'')^2]^N}-\frac{1}{[(x-x'')^2+(t-t'')^2]^N},$$

may serve as a barrier, where

$$x''=x_2+\frac{\rho}{\sqrt{1+k_2^2}},$$

and

$$t''=t_2+\frac{k_2\rho}{\sqrt{1+k_2^2}}.$$

It is obvious that conditions (1) and (2) of Lemma 2 are satisfied for the functions v_A. To show that condition (3) is satisfied, we need to use Taylor's formula at the point (t, x) and the fact that $\frac{\partial^2 v_A}{\partial x^2}-\frac{\partial v_A}{\partial t}<0$ in a sufficiently small

neighbourhood of the point A if N is chosen sufficiently great.

7. We have shown that an arbitrary infinite set of functions u^n contains a sequence that converges in the region G to the solution of the first boundary-value problem if the boundary of the region G satisfies the conditions listed in the preceding theorem. It is easy to show now that every sequence $\{u^n\}$ converges in the region G to the solution $U(t, x)$ of the first boundary-value problem.

This is true because, otherwise, there would be an infinite set of functions u^n, a point $(\bar{t}, \bar{x})$ belonging to G, and an $\varepsilon > 0$ such that, for every function u^n belonging to the set,

$$|u^n(\bar{t}, \bar{x}) - U(\bar{t}, \bar{x})| > \varepsilon.$$

This contradicts the fact that every infinite set of functions u^n contains a sub-sequence that converges in G to the solution of the first boundary-value problem, which, as we have shown, is unique.

Remarks 1: What we have said in this section is also applicable to the case of the heat-flow equation with an arbitrary number of independent variables.

2: The existence of a solution to the Dirichlet problem for Laplace's equation can be proven by the method of nets by a procedure analogous to that expounded in the present section*.

43. COMMENTS ON THE NETS METHOD

The method of nets or, as it is often called, the method of finite differences, is the most widely used method of approximate solution of partial differential equations. It has undergone an especially extensive development in the last few years in connection with its application to numerical calculation of fast-acting electronic computing machines.

We have already given several examples of the application of this method. In Section 10, we gave a brief description of the finite-difference method of approximate solution of the Cauchy problem for hyperbolic systems. In Section 16, we referred to the application of the method to numerical solution of the Cauchy problem for the wave equation. In Section 36, we applied the method to getting an approximate

* PETROVSKII, I.G., *Uspekhi matem. nauk*, 8, 161-170 (1941).

solution of the Dirichlet problem for Laplace's equation.

The method of nets has value in theory as well as application. By means of it, we can prove the existence of a solution to various boundary-value problems and we can also investigate the properties of the solutions. The existence of a solution to the first boundary-value problem for the heat-flow equation was proven by this method in the preceding section.

In the present section, we shall expound certain basic concepts associated with the method of nets. For simplicity of exposition, we shall consider only the case of two independent variables t and x and we shall confine ourselves to simple boundary-value problems for linear partial differential equations. We shall examine either the Cauchy problem or the problem with initial and boundary conditions.

1. The fundamental idea behind the method of nets consists in replacing the differential equation in the initial and boundary conditions with a system of finite-difference (algebraic) equations that approximately represent the original boundary-value problem.

To do this, we construct a net (that is, a finite or a countable set of points depending on one or several parameters) in the region G of the tx-plane in which we are seeking a solution. The points belonging to the net are called its nodes. Most frequently, one uses a rectangular net. The nodes of such a net have coordinates of the form $(t_0 + n\Delta t, x_0 + m\Delta x)$, where (t_0, x_0) denotes some point on the tx-plane and Δt and Δx are positive parameters called the steps of the net with respect to t and x respectively; n and m assume integral values. (An example of a nonrectangular net was given in Section 10 for the case of a hyperbolic system consisting of two equations. In this case, the net consists of the points of intersection of the tangents to the characteristics.)

We shall assume that the region G in which we wish to find the solution $u(t, x)$ is either a strip $0 < t < T$ or a rectangle $0 < t < T, 0 < x < 1$. In both cases, we shall use a rectangular net. In the first case, we set $t_0 = 0$. In the second case, we take $t_0 = x_0 = 0$, and $\Delta x = \frac{1}{M}$, where M is a positive integer.

For brevity in writing, we introduce the following notations:

$$\tau = \Delta t, \quad h = \Delta x, \quad u_m^n = u(n\tau, mh).$$

The set of all points of the net with the same value of n will

be called the layer with number n. For simplicity, we shall assume that τ depends on h (here, $\lim_{h \to 0} \tau = 0$) so that the rectangular net that we have constructed will be determined only by the parameter h.

There are various ways of constructing finite-difference equations that give approximate representations of a partial differential equation. The simplest method consists in replacing each of the partial derivatives appearing in the differential equation with a linear combination of values of $u(t, x)$ at the nodes of the net that approach the corresponding derivative as $h \to 0$. We shall call this linear combination a difference approximation of the corresponding derivative.

Let us give some examples. The derivative $\frac{\partial u}{\partial x}$ can be replaced at the point $t = n\tau$, $x = mh$ with any one of the following expressions:

$$\frac{\partial u}{\partial x} \approx \frac{u(t, x+h) - u(t, x)}{h} = \frac{u_{m+1}^n - u_m^n}{h}; \tag{43.1}$$

$$\frac{\partial u}{\partial x} \approx \frac{u(t, x) - u(t, x-h)}{h} = \frac{u_m^n - u_{m-1}^n}{h}; \tag{43.2}$$

$$\frac{\partial u}{\partial x} \approx \frac{u(t, x+h) - u(t, x-h)}{2h} = \frac{u_{m+1}^n - u_{m-1}^n}{2h}. \tag{43.3}$$

Let us estimate the errors in these approximating equations. By using Taylor's formula, we get

$$\frac{u(t, x+h) - u(t, x)}{h} = \frac{\partial u(t, x)}{\partial x} + \frac{h}{2} \frac{\partial^2 u(t, x+\theta_1 h)}{\partial x^2};$$

$$\frac{u(t, x) - u(t, x-h)}{h} = \frac{\partial u(t, x)}{\partial x} - \frac{h}{2} \frac{\partial^2 u(t, x-\theta_2 h)}{\partial x^2};$$

$$\frac{u(t, x+h) - u(t, x-h)}{2h} = \frac{\partial u(t, x)}{\partial x} + \frac{h^2}{6} \frac{\partial^3 u(t, x+\theta_3 h)}{\partial x^3};$$

Here, $\quad 0 < \theta_1 < 1, \quad 0 < \theta_2 < 1, \quad |\theta_3| < 1.$

The difference between any derivative and its difference approximation is called the approximation error or the remainder term. If the approximation error for some function is $O(h^k)$ we say that the order of approximation for this function is k.

For functions with bounded derivatives in the remainder

terms, the order of approximation of the approximating formulae (43.1) and (43.2) is equal to unity but the order of approximation of formula (43.3) is 2. One can construct approximating formulae for $\frac{\partial u}{\partial x}$ with as high an order of approximation as is desired. These formulae have a more complicated form since they contain the values of the function $u(t, x)$ at many neighbouring nodes of the net.

The difference approximations for $\frac{\partial u}{\partial t}$ are constructed in an analogous manner. When we replace the higher-order derivatives with difference approximations, we may treat a higher-order derivative as the result of successive differentiation of $u(t, x)$ with respect to t and x and we may successively apply the corresponding difference operations replacing differentiation.

For example,

$$\frac{\partial^2 u(t, x)}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial u(t, x)}{\partial x}\right) \approx \frac{\frac{\partial u(t, x+h)}{\partial x} - \frac{\partial u(t, x)}{\partial x}}{h}$$

$$\approx \frac{1}{h}\left[\frac{u(t, x+h) - u(t, x)}{h} - \frac{u(t, x) - u(t, x-h)}{h}\right] =$$

$$= \frac{u(t, x+h) - 2u(t, x) + u(t, x-h)}{h^2}.$$

Thus, we have obtained the approximating formula

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u(t, x+h) - 2u(t, x) + u(t, x-h)}{h^2}. \tag{43.4}$$

The approximation error for this formula is of the form

$$\frac{h^2}{12}\frac{\partial^4 u(t, x+\theta h)}{\partial x^4} = O(h^2), \text{where } |\theta| < 1.$$

Let us write our linear differential equation (which we wish to replace with a finite-difference equation) in the form

$$L(u) = f, \tag{43.5}$$

where $f = f(t, x)$ is a known function and $L(u)$ is a linear combination of the unknown function $u(t, x)$ and its partial derivatives. When we have expressed the derivatives occurring in $L(u)$ at the node (t, x) in terms of the corresponding finite-difference approximations and remainder terms, we obtain the following equation

$$\bar{L}_h(u) = f + \alpha_h. \tag{43.6}$$

Here $\overline{L}_h(u)$ is some finite-difference expression, that is, a linear combination of the values of u at the nodes of the net and α_h is the approximation error for the differential equation (43.5). This approximation error is given by the formula

$$\alpha_h = \overline{L}_h(u) - L(u). \tag{43.7}$$

If, on the right side of (43.6), we disregard the error of the approximation α_h, we obtain the approximating finite-difference equation

$$\overline{L}_h(\bar{u}) = f. \tag{43.8}$$

(Here and in what follows, we denote by $\bar{u}$ the solution of the finite-difference equation.)

The method of undetermined coefficients is also frequently used to construct approximating finite-difference equations. In this method, the entire left member $L(u)$ rather than the individual derivatives appearing in the differential equation (43.5) are approximated. To this end, we form a linear combination of the values of u at some set of nodes with undetermined coefficients. Then, all these values of u given by Taylor's formula are expressed in terms of the values of the function u and its derivatives at a single point (t, x) which may not even belong to the net. We obtain an expression that is linear with respect to the function $u(t, x)$ and its derivatives. The undetermined coefficients are then chosen in such a way that the expression obtained differs from $L(u)$ at the point (t, x) only by terms that approach zero as $h \longrightarrow 0$.

Let us take an example. Let us approximate the equation

$$L(u) \equiv \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = 0 \tag{43.9}$$

with the values of u at the nodes

$$(n\tau, mh), \; (n\tau, (m+1)h), \; ((n+1)\tau, mh), \; ((n+1)\tau, (m+1)h).$$

We place the point (t, x) at the centre of the rectangle whose corners are these nodes. For simplicity in calculation, we translate the coordinate origin to the point (t, x). We set

$$\overline{L}_h(u) = au\left(-\frac{\tau}{2}, -\frac{h}{2}\right) + bu\left(-\frac{\tau}{2}, \frac{h}{2}\right) + cu\left(\frac{\tau}{2}, \frac{h}{2}\right) + du\left(\frac{\tau}{2}, -\frac{h}{2}\right),$$

where a, b, c and d are undetermined coefficients.

From Taylor's formula, we have

$$\begin{aligned}\bar{L}_h(u) &= (a+b+c+d)\,u(0,0) + \frac{\tau}{2}(-a-b+c+d)\frac{\partial u}{\partial t}(0,0)\\ &+ \frac{h}{2}(-a+b+c-d)\frac{\partial u}{\partial x}(0,0) + \frac{\tau^2}{8}(a+b+c+d)\frac{\partial^2 u}{\partial t^2}(0,0)\\ &+ \frac{\tau h}{4}(a-b+c-d)\frac{\partial^2 u}{\partial t\,\partial x}(0,0)\\ &+ \frac{h^2}{8}(a+b+c+d)\frac{\partial^2 u}{\partial x^2}(0,0) + \ldots\end{aligned}$$

(The suspension points indicate terms of higher orders with respect to τ and h.) If we equate $\bar{L}_h(u)$ with $L(u)$ and then equate the coefficients of the corresponding derivatives, we obtain the following three equations:

$$a+b+c+d=0; \quad -a-b+c+d=\frac{2}{\tau};$$
$$-a+b+c-d=-\frac{2}{h}. \tag{43.10}$$

We supplement equations (43.10) with one more equation:

$$a-b+c-d=0, \tag{43.11}$$

and then determine a, b, c and d as the solution of the system of algebraic equations (43.10)-(43.11). Then, the expression $\bar{L}_h(u)$ will coincide with $L(u)$ with accuracy up to the terms represented by the suspension points, which tend to 0 as $h \longrightarrow 0$.

Returning to the original coordinate system, we obtain the following difference equation, which represents equation (43.9) approximately:

$$\bar{L}_h(\bar{u}) = \frac{(\bar{u}_{m+1}^{n+1} + \bar{u}_m^{n+1}) - (\bar{u}_{m+1}^n + \bar{u}_m^n)}{2\tau} - \frac{(\bar{u}_{m+1}^{n+1} + \bar{u}_{m+1}^n) - (\bar{u}_m^{n+1} + \bar{u}_m^n)}{2h} = 0. \tag{43.12}$$

We note that we might arrive at equation (43.12) by a different procedure. Specifically, the first term on the left side of this equation can be represented as follows:

$$\frac{1}{2}\left(\frac{u_{m+1}^{n+1} - u_{m+1}^n}{\tau} + \frac{u_m^{n+1} - u_m^n}{\tau}\right).$$

In this expression, the first term approximates the value

of $\frac{\partial u}{\partial t}$ at the point

$$\left(\left(n+\frac{1}{2}\right)\tau,\ (m+1)h\right);$$

with an error of $O(\tau^2)$. The second term approximates

$$\frac{\partial u}{\partial t}\left\{\left(n+\frac{1}{2}\right)\tau,\ mh\right\}.$$

with an error of the same order with respect to τ. It is easy to verify that the entire expression approximates $\frac{\partial u}{\partial t}$ at the point

$$\left\{\left(n+\frac{1}{2}\right)\tau,\left(m+\frac{1}{2}\right)h\right\}$$

with an error of

$$O(\tau^2)+O(h^2).$$

Analogously, we may show that the second term on the left side of (43.12) approximates $-\frac{\partial u}{\partial x}$ at the same point

$$\left\{\left(n+\frac{1}{2}\right)\tau,\ \left(m+\frac{1}{2}\right)h\right\}.$$

with the same error. Consequently, the left side of (43.12) approximates the left side of (43.9) with an error of $O(\tau^2)+O(h^2)$.

Let us give some examples of various finite-difference approximations for the left members of certain partial differential equations. The expressions in the parentheses indicate the order of the approximation error α_h with respect to τ and h;

1) $$L(u)\equiv\frac{\partial u}{\partial t}-\frac{\partial u}{\partial x}=0; \qquad (43.13)$$

$$\bar{L}_h(u)\equiv\frac{u_m^{n+1}-u_m^n}{\tau}-\frac{u_m^n-u_{m-1}^n}{h}\quad [\alpha_h=O(\tau)+O(h)]; \qquad (43.14)$$

$$\bar{L}_h(u)\equiv\frac{u_m^{n+1}-u_m^n}{\tau}-\frac{u_{m+1}^n-u_m^n}{h}\quad [\alpha_h=O(\tau)+O(h)]; \qquad (43.15)$$

$$\bar{L}_h(u)\equiv\frac{u_m^{n+1}-u_m^{n-1}}{2\tau}-\frac{u_{m+1}^n-u_{m-1}^n}{2h}$$

$$[\alpha_h=O(\tau^2)+O(h^2)]; \qquad (43.16)$$

$$\bar{L}_h(u) \equiv \frac{(u_{m+1}^{n+1} - u_{m+1}^n) + (u_m^{n+1} - u_m^n)}{2\tau}$$

$$- \frac{(u_{m+1}^{n+1} - u_m^{n+1}) + (u_{m+1}^n - u_m^n)}{2h} \quad [\alpha_h = O(\tau^2) + O(h^2)]. \quad (43.17)$$

2) $$L(u) \equiv \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0; \quad (43.18)$$

$$\bar{L}_h(u) \equiv \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} \quad [\alpha_h = O(\tau) + O(h^2)]; \quad (43.19)$$

$$\bar{L}_h(u) \equiv \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2} \quad (43.20)$$

$$[\alpha_h = O(\tau) + O(h^2)];$$

$$\bar{L}_h(u) \equiv \frac{u_m^{n+1} - u_m^n}{\tau} - \frac{1}{2}\left[\frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2} + \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2}\right] \quad (43.21)$$

$$[\alpha_h = O(\tau^2) + O(h^2)];$$

$$\bar{L}_h(u) \equiv \frac{3}{2}\frac{u_m^{n+1} - u_m^n}{\tau} - \frac{1}{2}\frac{u_m^n - u_m^{n-1}}{\tau} - \frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2} \quad (43.22)$$

$$[\alpha_h = O(\tau^2) + O(h^2)];$$

$$\bar{L}_h(u) \equiv \frac{u_m^{n+1} - u_m^{n-1}}{2\tau} - \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} \quad (43.23)$$

$$[\alpha_h = O(\tau^2) + O(h^2)].$$

3) $$L(u) \equiv \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0; \quad (43.24)$$

$$\bar{L}_h(u) \equiv \frac{u_m^{n+1} - 2u_m^n + u_m^{n-1}}{\tau^2} - \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} \quad (43.25)$$

$$[\alpha_h = O(\tau^2) + O(h^2)];$$

$$\bar{L}_h(u) \equiv \frac{u_m^{n+1} - 2u_m^n + u_m^{n-1}}{\tau^2} - \frac{1}{2}\left(\frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2} + \frac{u_{m+1}^{n-1} - 2u_m^{n-1} + u_{m-1}^{n-1}}{h^2}\right) \quad [\alpha_h = O(\tau^2) + O(h^2)]. \quad (43.26)$$

All these expressions can be obtained from the left members of the corresponding differential equations either by replacing the derivatives according to formulae of the form

(43.1)-(43.4) (and analogous formulae for derivatives with respect to t) or by the method of undetermined coefficients.

In the examples that we examined above, the approximation error was estimated in absolute value. Here, we assumed that the solution of the differential equation is sufficiently smooth as to admit a representation by Taylor's formulae up to terms of the necessary order and that the derivatives appearing in the remainder term are bounded. If the solution is less smooth than this, the approximation error may have a lower order.

To obtain a difference boundary-value problem, we need to replace not only the differential equation but also the initial and boundary conditions with difference equations. If the derivatives of the solution do not appear in the initial and boundary conditions, with the above-indicated choice of a rectangular net, the given boundary conditions determine directly the values of the unknown function at the corresponding boundary nodes of the net. However, if the boundary conditions contain derivatives of the unknown function, these derivatives can be replaced with difference approximations as indicated above. We then obtain difference boundary conditions approximating with some error the boundary conditions for the differential equation. There are also other ways of approximating boundary conditions.

We give an example. Suppose that we need to replace the boundary condition

$$l(u) \equiv A(t)\,u(t,\ 0) + B(t)\frac{\partial u}{\partial x}(t,\ 0) = C(t).$$

with difference equations. The simplest method consists in replacing $\frac{\partial u}{\partial x}(t,\ 0)$ with the expression

$$\frac{u(t,\ h) - u(t,\ 0)}{h}.$$

We then obtain the difference boundary condition

$$\bar{l}_h(\bar{u}) \equiv A(n\tau)\,\bar{u}_0^n + B(n\tau)\frac{\bar{u}_1^n - \bar{u}_0^n}{h} = C(n\tau).$$

The approximation error, that is, the difference $l(u) - \bar{l}_h(u)$, is $O(h)$. If we replace $\frac{\partial u}{\partial x}(t,\ 0)$ with the expression

$$\frac{3u_0^n - 4u_1^n + u_2^n}{h},$$

we obtain the more exact approximation

$$l(u) - \bar{l}_h(u) = O(h^2).$$

Let us see what values the indices m and n assume in the difference equation approximating the differential equation and the difference equations approximating the initial and boundary conditions. The system of algebraic equations that we obtain is called the difference scheme of the corresponding problem for the partial differential equation.

The difference schemes for problems of the type that we have been considering differ in the numbers of layers appearing in the difference equation approximating the differential equation. We speak of two-layer, three-layer, etc. schemes. Two-layer schemes are obtained, for example, with the aid of the difference expressions (43.14), (43.15), (43.17), (43.19), (43.20), and (43.21). Formulae (43.16), (43.22), (43.23), (43.25), and (43.26) lead to three-layer schemes. In solving problems with initial conditions by means of a two-layer scheme, it is sufficient to give the values of $\bar{u}$ at the nodes of one initial layer ($n = 0$). In the case of a three-layer scheme, it is necessary to give the values of $\bar{u}$ at the nodes of two adjacent initial layers ($n = 0$ and $n = 1$).

If the values of the unknown function in the $(n+1)$st layer are expressed directly from the difference equations of the scheme in terms of the values of this function on the preceding layers, the difference scheme is said to be explicit. If instead it is necessary to solve a system of equations, the scheme is said to be implicit. Formulae (43.17), (43.20), (43.21), (43.22) and (43.26) lead to implicit schemes*.

We introduce certain notations. We shall write the initial and boundary conditions for the equation $L(u) = f$ in the form

$$l_i(u) = \varphi_i, \quad i = 1, 2, \ldots, p \quad (\text{intitial conditions}); \qquad (43.27)_1$$

$$l_i(u) = \varphi_i, \quad i = p+1, p+2, \ldots, s \quad (\text{boundary conditions}). \qquad (43.27)_2$$

Here, the φ_i are unknown functions given on certain portions of the boundary of the region G and the $l_i(u)$ are linear combinations of the unknown function and its partial derivatives. We shall write the initial and boundary conditions

* A convenient method for solving systems of equations that arise from use of implicit difference schemes is expounded in the book by BEREZIN, I.S. and ZHIDKOV, N.P., *Computational methods*, Fizmatgiz, II, Chapter X, Section 6 (1959).

for the difference scheme in the form

$$\bar{l}_{ih}(\bar{u}) = \varphi_{ih}, \quad i = 1, 2, \ldots, p \ (\text{ initial conditions }); \quad (43.28)_1$$

$$\bar{l}_{ih}(\bar{u}) = \varphi_{ih}, \quad i = p+1, p+2, \ldots, s \ (\text{ boundary conditions }). \quad (43.28)_2$$

The values of the approximate solution at the nodes of the net, that is, the quantities $\bar{u}_m^n$ can be regarded as the components of a vector in some linear space the dimension of which is determined by the number of nodes in the net. We shall denote this vector by $\bar{u}_h$.

The right sides of the difference equations approximating the differential equation and the right sides of the difference initial and boundary conditions also constitute a vector, which we shall denote by F_h. We denote by R_h the matrix of the coefficients of the system of equations constituting the difference scheme. In these notations, the difference scheme, that is, the set of equations (43.8)-(43.28), can be written

$$R_h \bar{u}_h = F_h. \tag{43.29}$$

The difference $R_h u - F_h \equiv r_h$ is called the approximation error of the difference scheme (43.29) for the given solution $u(t, x)$ of the boundary-value problem (43.5)-(43.27).

From the point of view of applications, the estimate of the difference $u - \bar{u}_h$ between the exact and approximate solutions is of fundamental significance. This difference is defined only at the nodes of the net. To obtain an estimate for r_h and $u - \bar{u}_h$, it is natural to use the concept of a norm in a linear space. What is actually used in the estimates of the approximation error of r_h that were made above is the norm defined for the element $w = \{w_m^n\}$ by the formula

$$\| w \|_C = \sup_{n,\, m} | w_m^n |. \tag{43.30}$$

However, we need to consider other norms as well, specifically, the so-called integral norms. Examples of such norms are given in subsection 5. Similar norms are used in particular, in the cases in which the approximation error does not approach zero in absolute value as $h \to 0$ and even increases without bound at individual points but is small in some integral sense, for example, in mean.

Also, different norms may be used to measure the difference $u - \bar{u}_h$.

Let us assume that several norms are given for r_h and

for $u - \bar{u}_h$. (These norms, in general, do not coincide.)

The following are two basic definitions:

I. The difference scheme (43.29) approximate the boundary-value problem (43.5)-(43.27) for a given solution $u(t, x)$ of that boundary-value problem if the approximation error r_h approaches zero as $h \to 0$. If $r_h = O(h^k)$, we say that the order of approximation is equal to k.

II. The difference scheme (43.29) is said to be convergent for a given solution $u(t, x)$ of the boundary-value problem (43.5)-(43.27) if the difference $u - \bar{u}_h$ approaches zero as $h \to 0$. If $u - \bar{u}_h = O(h^q)$ we say that the order of convergence is equal to q.

With these definitions, it is assumed that r_h and $u - \bar{u}_h$ approach zero in terms of the corresponding norms. It is in this sense that we should understand the equations $r_h = O(h^k)$ and $u - \bar{u}_h = O(h^q)$ (Thus, for example, the equation $r_h = O(h^k)$ means that $\| r_h \| h^{-k} \leq C_1 = \text{const}$).

2. Certain simple examples show that not every approximating difference scheme converges even for arbitrary smoothness of the exact solution.

Consider the following Cauchy problem:

$$L(u) \equiv \frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} = 0;$$

$$u(0, x) = \varphi(x), \quad -\infty < x < \infty. \tag{43.31}$$

Let us suppose that $\varphi(x)$ has a bounded second derivative. Then, the exact solution $u(t, x) = \varphi(x + t)$ of this problem has bounded second partial derivatives. Therefore, the difference scheme

$$\left.\begin{aligned} &\bar{L}_h(\bar{u}_h) \equiv \frac{\bar{u}_m^{n+1} - \bar{u}_m^n}{\tau} - \frac{\bar{u}_m^n - \bar{u}_{m-1}^n}{h} = 0, \\ &\bar{u}_m^0 = \varphi(mh) \\ &\left(n = 0, 1, \ldots, \left[\frac{T}{\tau}\right] - 1; m = 0, \pm 1, \ldots\right) \end{aligned}\right\} \tag{43.32}$$

approximates the problem (43.31) with an error $O(\tau) + O(h)$ (in the sense of the norm (43.30)).

Let us show that the difference scheme (43.32) cannot be convergent no matter what the relationship is between τ and h. It follows from equations (43.32) that the value of

the function $\bar{u}_h$ at the node $(t+\tau, x)$ is defined in terms of its values at the nodes $(t, x-h)$ and (t, x). Let us consider some particular node (t_0, x_0). Let us consider some particular node (t_0, x_0) It is easy to see that the value of $\bar{u}_h(t_0, x_0)$ that is found by use of (43.32) is uniquely defined in the terms of the value of the function $\varphi(x)$ for $x \leqslant x_0$. On the other hand, the value of the exact solution $u(t, x)$ at the point (t_0, x_0) is equal to $\varphi(x_0+t_0)$; that is, it is defined by the value of $\varphi(x)$ for $x > x_0$.

Suppose that $\varphi(x) \equiv 0$ for $x \leqslant x_0$ and that $\varphi(x) > 0$ for $x > x_0$. Then, $\bar{u}_h(t_0, x) = 0$ for all $x \leqslant x_0$ However,

$$u(t_0, x) = \varphi(x + t_0) > 0$$

for $x_0 - t_0 < x < \infty$. Thus, in the region

$$\{x_0 - t < x < x_0, \ 0 < t < \infty\}$$

the difference $u - \bar{u}_h$ between the exact solution of the Cauchy problem (43.31) and the solution of the difference equations (43.32) is positive and independent of h. Consequently, this difference cannot approach zero as $h \to 0$.

Let us now look at the difference scheme

$$\bar{L}_h(\bar{u}_h) \equiv \frac{\bar{u}_m^{n+1} - \bar{u}_m^n}{\tau} - \frac{\bar{u}_{m+1}^n - \bar{u}_m^n}{h} = 0; \qquad (43.33)_1$$

$$\bar{u}_m^0 = \varphi(mh); \qquad (43.33)_2$$

$$\left(n = 0, 1, \ldots, \left[\frac{T}{\tau}\right] - 1; \ m = 0, \pm 1, \ldots\right).$$

which is outwardly very similar to (43.32). Let us show that this scheme is convergent (in the sense of the norm (43.30)) for $\frac{\tau}{h} = c \leqslant 1$ (for $c \equiv$ const)

Suppose that $\alpha_h = \bar{L}_h(u) - L(u)$. Since $L(u) = 0$, we have

$$\bar{L}_h(u) = \alpha_h. \qquad (43.34)$$

As a result of the estimate (43.15) for α_h and the condition $\tau/h = \text{const}$, we have $\alpha_h = O(h)$; that is, α_h converges uniformly to zero as $h \to 0$.

By using (43.33) and (43.34), we obtain for the difference $u - \bar{u}_h = v_h$ the equation

$$\bar{L}_h(v_h) = \alpha_h \qquad (43.35)$$

with the initial condition $v_m^0 = 0$. When we solve equation (43.35) for v_m^{n+1}, we obtain

$$v_m^{n+1} = \tau\alpha_m^n + (1-c)\,v_m^n + cv_{m+1}^n. \tag{43.36}$$

Suppose that

$$V_n = \sup_{-\infty<m<\infty} |v_m^n|,$$

and

$$A_h = \sup_{n,\,m} |\alpha_m^n| = O(h).$$

From (43.36), we obtain

$$V_{n+1} \leqslant \tau A_h + V_n \left(n = 0,\ 1, \ldots,\ \left[\frac{T}{\tau}\right] - 1\right).$$

When we sum these inequalities over n from 0 to $N-1$ (where $N\tau \leqslant T$ and take into account the fact that $V_0 \equiv 0$, we arrive at the relation

$$V_N \leqslant A_h T = O(h). \tag{43.37}$$

From this follows the uniform converge of $\bar{u}_h - u$ to zero in the strip $\{0 \leqslant t \leqslant T, -\infty < x < \infty\}$ and the estimate

$$u - \bar{u}_h = O(h).$$

We note that the equation $c = \frac{\tau}{h} \leqslant 1$ is essential for the convergence of the difference scheme (43.33). For $c > 1$ the scheme (43.33) will not converge, which can easily be shown in a manner analogous to what was done for the scheme (43.32). Thus, convergence of the scheme depends not only on the form of the difference equations but also on the choice of relationships between the steps of the net.

3. In proving the convergence of the difference scheme (43.33), we actually used only two properties of this soheme. First, since the scheme (43.33) is an approximating one, we see that the exact solution of the Cauchy problem (43.31) satisfies the difference equation (43.34) with the right member α_h approaching zero as $h \to 0$. Second, we used the fact that, with the same initial condition, the difference v_h between the solution of equation (43.34) and that of equation

$(43.33)_1$ approaches zero as $\alpha_h \to 0$. This fact, which indicates the continuous dependence of the solution of the difference equation (43.34) on the right member, was established with the aid of inequality (43.37). It follows from (43.37) that $|v_h| < \varepsilon$ if $|\alpha_h| < \delta$, where $\delta > 0$ depends on ε but not on h.

In the present example, the approximation error lies only in the difference equation approximating the differential equation. In other cases, the approximation error appears also in the difference equations that are approximate representations of the initial and boundary conditions. Therefore, improving convergence, we need to investigate the independence of the solution of the difference equation not only on the right member of equation (43.8) but also on the right members of the initial and boundary conditions (43.27).

In analog with the definition of a correctly stated boundary-value problem for a differential equation (see Section 8), we speak of a correct difference scheme.

The difference scheme (43.29) is said to be correct if the following conditions hold: (1) For every sufficiently small positive number h (where $0 < h < h_0$) and for an arbitrary right member F_h, there exists a unique solution $\bar{u}_h$ of equation (43.29) that depends continuously on F_h, and (2) this continuous dependence is uniform with respect to h. The second requirement means that, for an arbitrary $\varepsilon > 0$, there exists a $\delta > 0$, independent of h, such that for every change in F_h that does not exceed δ (according to some norm), the corresponding change in $\bar{u}_h$ will not exceed ε (in general, according to some other norm).

We may make the following general assertion from the definitions given above: for a suitable choice of norms in the definitions of approximation, convergence, and correctness, the solutions of a correct difference scheme approach the solution of the boundary-value problem for the differential equation as h approaches 0 if this solution exists.

We have proved this assertion for the example of the scheme (43.33). The general case is proven in an analogous manner. There, the boundary-value problem (43.5)-(43.27) may be either linear or nonlinear. In the case of a nonlinear boundary-value problem, it can be shown that the order of convergence is equal to the order of approximation. This enables us to justify the simple method (often used in practice) of estimating the error of an approximate solution by comparison of the approximate solutions $\bar{u}_h$ that are ob-

tained for different values of h. All these questions are expounded in detail in the book by V.S. Ryaben'kii and A.F. Filippov, Ob ustoichivosti raznostnykh uravnenii (The stability of difference equations), Gosekhizdat, 1956.

4. From the above definition of a correct difference scheme, the solution of such a scheme depends continuously (uniformly with respect to h) on the right member f of the difference equation (43.8) approximating the differential equation, on the right members φ_{ih} (for $i=1, \ldots, p$) of the initial conditions $(43.28)_1$, and on the right members φ_{ih} (for $i=p+1, \ldots, s$) of the boundary conditions $(43.28)_2$. When we are investigating the correctness of a linear difference scheme, it is sufficient for us to study separately the dependence of the solution of each of these three quantities. If the difference scheme has a solution for arbitrary values of φ_{ih} (for $i=1, \ldots, p$) that depends continuously (uniformly with respect to h) on φ_{ih} (for $i=1, \ldots, p$), the difference scheme is said to be stable with respect to the initial conditions. Stability with respect to the right member of the difference equation approximating the differential equation and stability with respect to the boundary conditions are defined analogously. Obviously, a linear difference scheme that is stable with respect to the initial condition, with the right member, and with respect to the boundary conditions is correct.

At the present time, sufficiently general methods of investigating the convergence of difference schemes have not been found. As a rule, it is easiest to investigate the stability with respect to the initial conditions. It can be shown that stability with respect to the initial conditions implies stability with respect to the right member for a rather broad class of schemes. The question of stability with respect to the boundary conditions has as yet received only slight study.

As examples, let us look at certain methods of investigating convergence with respect to initial conditions.

Suppose that $\delta\bar{u}=\{\delta\bar{u}_m^n\}$ is the change in the solution $\bar{u}_m^n$ of the linear difference scheme (43.8)-(43.28), caused by a change $\delta\varphi_{ih}$ (for $i=1, \ldots, p$) in the initial conditions $(43.28)_1$. It is easy to see that $\delta\bar{u}$ is a solution of the following difference scheme:

$$\bar{L}_h(\delta\bar{u})=0; \tag{43.38}$$

$$\begin{aligned} \bar{l}_{ih}(\delta\bar{u}) &= \delta\varphi_{ih},\ i=1,\ \dots,\ p\ (\text{ initial conditions });\\ \bar{l}_{ih}(\delta\bar{u}) &= 0,\ i=p+1,\ \dots,\ s\ (\text{boundary conditions}). \end{aligned} \tag{43.38}$$

Therefore, in investingating the stability on the basis of the initial conditions, we may confine ourselves to a study of schemes of the form (43.38). For brevity in notation, we shall henceforth write $\bar{u}_m^n$ instead of $\delta\bar{u}_m^n$.

Let us look at the scheme (43.33) again. Inequality (43.37) means that this scheme is stable at $c=\frac{\tau}{h}\leqslant 1$ with respect to the right member. Let us show that this scheme is also stable for $c\leqslant 1$ with respect to the initial conditions if we use the norms

$$\|\bar{u}_h\| = \sup_{n,\,m} |\bar{u}_m^n|, \quad \|\varphi\| = \sup_m |\varphi(mh)|.$$

We set

$$U_n = \sup_m |\bar{u}_m^n|.$$

From (43.33), we have

$$\bar{u}_m^{n+1} = (1-c)\,\bar{u}_m^n + c\bar{u}_{m+1}^n.$$

From this, we get $U_{n+1}\leqslant U_n$ and consequently

$$U_n \leqslant U_0 \quad \left(n=0,\ 1,\ \dots,\ \left[\frac{T}{\tau}\right]\right).$$

for $c\leqslant 1$ .Therefore, $\|\bar{u}_h\|\leqslant\varepsilon$ if $|\varphi\|\leqslant\varepsilon$. This proves the stability with respect to the initial conditions for $c\leqslant 1$.

Let us now show that if $c=1+\mu$ where $\mu>0$, the scheme (43.33) is unstable with respect to the initial conditions. Suppose that $\bar{u}_m^0=(-1)^m\varepsilon$. It is easy to verify that the solution in this case is of the form

$$\bar{u}_m^n = (-1)^{n+m}(1+2\mu)^n\varepsilon.$$

For every fixed $t=n\tau$ as $h\to 0$ the solution increases without bound, and, what is more, faster than any power of $1/h$ since

$$|\bar{u}_m^n| = \varepsilon(1+2\mu)^n = \varepsilon e^{\frac{Kt}{h}}, \text{where } K = \frac{\ln(1+2\mu)}{1+\mu}.$$

As a second example, let us consider the Cauchy problem for the heat-flow equation (43.18). We construct a difference

scheme in accordance with (43.19). Let us set $\frac{\tau}{h^2}=c$. When we solve the corresponding difference equation $\bar{L}_h(\bar{u}_h)=0$ for $\bar{u}_m^{n+1}$, we obtain

$$\bar{u}_m^{n+1}=c\bar{u}_{m+1}^n+(1-2c)\bar{u}_m^n+c\bar{u}_{m-1}^n. \tag{43.39}$$

If $c\leqslant\frac{1}{2}$, all the coefficients on the right side of (43.39) are nonnegative and their sum is equal to unity. Just as in the preceding example, it follows from this that the least upper bound of the absolute value of the solution does not increase when we go from n to $n+1$. Thus, the difference scheme constructed in accordance with (43.19) is stable with respect to the initial conditions when $c\leqslant\frac{1}{2}$.

For $c=\frac{1}{2}+\mu$, where $\mu>0$, this difference scheme is unstable with respect to the initial conditions. To prove this assertion, we again set $\bar{u}_m^0=(-1)^m\varepsilon$; after some simple calculations, we obtain

$$|\bar{u}_m^n|=\varepsilon(4c-1)^n=\varepsilon e^{\frac{Kt}{h^2}}, \tag{43.40}$$

$$\text{where } K=\frac{2\ln(1+4\mu)}{1+2\mu},\quad t=n\tau.$$

Proof of the stability of the difference schemes with respect to the initial conditions in these examples was based exclusively on the fact that the sum of the absolute values of the coefficients in the formulae expressing $\bar{u}_m^{n+1}$ in terms of the values of the solution at the nodes of the nth layer did not exceed unity. This sum of the absolute values of the coefficients is called the index of the difference scheme. For a scheme to be stable with respect to the initial conditions, it is sufficient that the index of the scheme not exceed $1+C\tau$, where C is some constant. This is true because under this condition, for arbitrary $t=n\tau\leqslant T$

$$\sup_m|\bar{u}_m^n|\leqslant(1+C\tau)^{\frac{T}{\tau}}\sup_m|\bar{u}_m^0|\leqslant e^{CT}\sup_m|\bar{u}_m^0|,$$

from which it follows that the scheme is stable with respect to the initial conditions.

Sometimes, we can investigate the stability with respect to the initial conditions by using properties analogous to the

maximum principle for solutions of the heat-flow equation. As an example, let us consider the difference scheme

$$\frac{\bar{u}_m^{n+1} - \bar{u}_m^n}{\tau} = \frac{\bar{u}_{m+1}^{n+1} - 2\bar{u}_m^{n+1} + \bar{u}_{m-1}^{n+1}}{h^2},$$

$$\bar{u}_0^{n+1} = \bar{u}_M^{n+1} = 0, \quad \bar{u}_m^0 = \varphi(mh)$$

$$\left(n = 0, 1, \ldots, \left[\frac{T}{\tau}\right] - 1; \; m = 1, 2, \ldots, M-1\right),$$

which approximates the first boundary-value problem for the heat-flow equation in the rectangle $\{0 < t < T, 0 < x < 1\}$ with the conditions

$$u(0, x) = \varphi(x); \; u(t, 0) = u(t, 1) = 0. \qquad (43.41)$$

From the lemma of paragraph 2 of Section 42, we have

$$\sup_{n, m} |\bar{u}_m^n| \leqslant \sup_m |\bar{u}_m^0|$$

for an arbitrary value of $\frac{\tau}{h^2}$. (Proof of this lemma does not depend on the value of $\frac{\tau}{h^2}$). From this it follows that the scheme in question is stable with respect to the initial conditions for an arbitrary value of $\frac{\tau}{h^2}$.

5. In addition to particular techniques that are based primarily on special properties of some difference scheme or other, there are two general methods that are used in investigating stability with respect to initial conditions, namely, the method of separation of variables (for boundary-value problems with initial and boundary conditions) and the Fourier-integral method (for the Cauchy problem).

Let us give some examples illustrating the method of separation of variables.

Consider the difference scheme

$$\frac{\bar{u}_m^{n+1} - \bar{u}_m^n}{\tau} = \frac{1}{2}\left(\frac{\bar{u}_{m+1}^{n+1} - 2\bar{u}_m^{n+1} + \bar{u}_{m-1}^{n+1}}{h^2} + \frac{\bar{u}_{m+1}^n - 2\bar{u}_m^n + \bar{u}_{m-1}^n}{h^2}\right), \qquad (43.42)_1$$

$$\bar{u}_0^n = 0, \; \bar{u}_M^n = 0, \; \bar{u}_m^0 = \varphi(mh) \qquad (43.42)_2$$

$$\left(n=0,\ 1,\ \ldots,\ \left[\frac{T}{\tau}\right]-1;\ m=1,\ 2,\ \ldots,\ M-1\right),$$

approximating the first boundary-value problem for the heat-flow equation (43.18) with conditions (43.41). In analogy with the method of separation of variables for a differential equation, we first seek solutions of equation $(43.42)_1$ satisfying zero boundary conditions and possessing the particular form

$$\bar{u}_m^n = T(n)\,X(m).$$

If we substitute this expression into $(43.42)_1$, we obtain after separating the variables

$$\frac{2h^2}{\tau}\,\frac{T(n+1)-T(n)}{T(n+1)+T(n)} = \frac{X(m+1)-2X(m)+X(m-1)}{X(m)} = \lambda, \quad (43.43)$$

where λ does not depend on n or m. To determine λ and $X(m)$, we have the following difference boundary-value problem, analogous to the Sturm-Liouville problem (see Section 20):

$$X(m+1) - 2X(m) + X(m-1) = \lambda X(m) \quad (43.44)$$
$$(m=1,\ 2,\ \ldots,\ M-1),$$

$$X(0) = X(M) = 0. \quad (43.45)$$

It is natural for us to call those values of λ at which a nontrivial solution to the problem (43.44)-(43.45) exists the eigenvalues of this problem and the nontrivial solutions $X(m)$ themselves the eigenfunctions.

Let us find the general solution of equation (43.44). In analogy with the familar method of solving ordinary linear differential equations with constant coefficients, we first seek particular solutions of (43.44) of the form

$$X(m) = e^{kx} = e^{kmh} = q^m, \text{where } q = e^{kh}.$$

To determine q from (43.44), we obtain the so-called oharacteristic equation

$$q^2 - (2+\lambda)\,q + 1 = 0. \quad (43.46)$$

Suppose that q_1 and q_2 are distinct roots of (43.46). Then, any solution of (43.44) can be represented in the form

$$X(m) = C_1 q_1^m + C_2 q_2^m,$$

where C_1 and C_2 are constants. This is true, because, as

can easily be verified, every function of this form satisfies equation (43.44). Furthermore, it follows immediately from (43.44) that an arbitrary solution $X(m)$ of this equation is uniquely determined if the values of $X(m)$ are given at two adjacent points: $X(m_0-1)=a$, and $X(m_0)=b$. But these last conditions can be satisfied if we take C_1 and C_2 such that

$$C_1 q_1^{m_0-1}+C_2 q_2^{m_0-1}=a,$$
$$C_1 q_1^{m_0}+C_2 q_2^{m_0}=b,$$

This system is obviously compatible for arbitrary a and b since $q_1 \neq q_2$.

By an analogous procedure, we may show that, for $q_1=q_2=q$, an arbitrary solution of equation (43.44) can be represented in the form

$$X(m)=(C_1+C_2 m)\, q^m.$$

It is not difficult to verify that if a function of this form satisfies the boundary conditions (43.45), it is identically equal to zero. Therefore, let us seek a solution to the problem (43.44)-(43.45) in the form

$$X(m)=C_1 q_1^m+C_2 q_2^m,$$

where $q_1 \neq q_2$.

We set $q_1=q$. Then, $q_2=q^{-1}$ since $q_1 q_2=1$. Consequently,

$$X(m)=C_1 q^m+C_2 q^{-m}$$

By using the boundary condition $X(0)=0$, we see that $C_2=-C_1$ and

$$X(m)=C_1(q^m-q^{-m}). \qquad (43.47)$$

The second boundary condition, $X(M)=0$, leads to the equation

$$q^{2M}=1,$$

so that

$$q=e^{i\frac{\pi k}{M}}, \quad k=0,\ 1,\ \ldots,\ 2M-1. \qquad (43.48)$$

If we set $C_1=\frac{1}{2i}$, we obtain from (43.47) and (43.48) the $M-1$ eigenfunctions

$$X_k(m)=\sin\frac{\pi k m}{M}, \quad k=1,\ 2,\ \ldots,\ M-1. \qquad (43.49)$$

(The remaining values of k indicated in (43.48) lead to the same eigenfunctions up to a factor of -1.) The corresponding eigenvalues can be found from the relation $\lambda = q + q^{-1} - 2$, which follows from equation (43.46) on the basis of Vieta's formula. We obtain

$$\lambda_k = -4 \sin^2 \frac{\pi k}{2M}, \quad k = 1,\ 2,\ \ldots,\ M-1. \quad (43.50)$$

Consider the eigenfunctions (43.49) at the nodes of the net $x = mh$, for $m = 1, 2, \ldots, M-1$). By virtue of the system (43.44)-(43.45), these are the eigen vectors of a real symmetric matrix made up of the coefficients of equations (43.44). Since all the eigenvalues (43.50) are distinct, the eigenfunctions (43.49) are linearly independent. They constitute a basis in the $(M-1)$-dimensional linear space composed of the functions $\{f(x)\}$, which we are considering only at the nodes of the net ($x = mh$, for $m = 1, 2, \ldots, M-1$).

We define a scalar product in this space by

$$(f,\ g)_h = h \sum_{m=1}^{M-1} f(mh)\, g(mh). \quad (43.51)$$

On the basis of a familar theorem in algebra*, the eigenfunctions (49.49) are mutually orthogonal in the sense of the scalar product (43.51); that is,

$$(X_k,\ X_l)_h = 0 \text{ when } k \neq l\ ).$$

It is easy to verify that $(X_k,\ X_k)_h = \frac{1}{2}$. Therefore, the system of functions

$$\tilde{X}_k(m) = \sqrt{2} \sin \frac{\pi k m}{M}, \quad k = 1,\ 2,\ \ldots,\ M-1, \quad (43.52)$$

constitutes an orthonormal basis in the space of functions on the net.

Let us now find the functions $T_k(n)$ (for $k = 1, 2, \ldots, M-1$). From equation (43.43), we obtain

$$T_k(n+1) = \frac{1 + \frac{\tau}{2h^2} \lambda_k}{1 - \frac{\tau}{2h^2} \lambda_k} T_k(n).$$

* See, for example, GEL'FAND, I.M., *Lectures on linear algebra*, Gostekhizdat, 121 (1951).

Therefore,

$$T_k(n) = A_k (s_k)^n, \tag{43.53}$$

where

$$s_k = \frac{1 + \frac{\tau}{2h^2}\lambda_k}{1 - \frac{\tau}{2h^2}\lambda_k}, \tag{43.54}$$

and A_k is an arbitrary constant. We note that $|s_k| < 1$ since $\lambda_k < 0$.

Following the basic idea of the method of separation of variables, let us seek a solution to the difference boundary-value problem (43.42) in the form

$$\bar{u}_m^n = \sum_{k=1}^{M-1} a_k (s_k)^n \tilde{X}_k(m), \tag{43.55}$$

where the a_k are constants to be chosen so as to satisfy the initial condition $\bar{u}_m^0 = \varphi(mh)$.

For this, we need to set $a_k = (\varphi, \tilde{X}_k)_h$.

From equation (43.55), we obtain

$$(\bar{u}^n, \bar{u}^n)_h = \sum_{k=1}^{M-1} |a_k|^2 |s_k|^{2n},$$

since the functions $\tilde{X}_k(m)$ form an orthonormal system. In particular, if we set $n = 0$, we get

$$(\bar{u}^0, \bar{u}^0)_h = (\varphi, \varphi)_h = \sum_{k=1}^{M-1} |a_k|^2.$$

Since $|s_k| < 1$, we have $(\bar{u}^n, \bar{u}^n)_h \leqslant (\varphi, \varphi)_h$ for $n > 0$. It follows from this that the solution is stable with respect to the initial conditions if the change in the initial conditions is measured in terms of the norm $\sqrt{(\varphi, \varphi)_h}$ and the change in the solution in terms of the norm $\sup\limits_n \sqrt{(\bar{u}^n, \bar{u}^n)_h}$.

We can investigate the following difference schemes for the same boundary-value problem (43.18)-(43.41) in an analogous manner by using the method of separation of variables:

$$\left.\begin{aligned} &1)\ \frac{\bar{u}_m^{n+1} - \bar{u}_m^n}{\tau} = \frac{\bar{u}_{m+1}^{n+1} - 2\bar{u}_m^{n+1} + \bar{u}_{m-1}^{n+1}}{h^2}, \\ &\bar{u}_0^n = \bar{u}_M^n = 0, \quad \bar{u}_m^0 = \varphi(mh). \end{aligned}\right\} \tag{43.56}$$

$$\left.\begin{aligned} &2)\ \frac{\bar{u}_m^{n+1}-\bar{u}_m^n}{\tau}=\frac{\bar{u}_{m+1}^n-2\bar{u}_m^n+\bar{u}_{m-1}^n}{h^2}, \\ &\bar{u}_0^n=\bar{u}_M^n=0, \quad \bar{u}_m^0=\varphi(mh). \end{aligned}\right\} \tag{43.57}$$

We suggest that the reader prove that the difference scheme (43.56) is stable with respect to the initial conditions for an arbitrary value of $\frac{\tau}{h^2}$ and that the scheme (43.57) is stable if $\frac{\tau}{h^2}\leqslant\frac{1}{2}$.

We have already investigated the difference scheme (43.56) with the aid of the maximum principle (with a different choice of norms) and we have shown that it is stable with respect to the initial conditions with no restrictions of any kind imposed on the value of $\frac{\tau}{h^2}$. We have also examined the difference equations in the scheme (43.57) for the Cauchy problem and we showed that the corresponding difference scheme is stable with respect to the initial conditions if $\frac{\tau}{h^2}\leqslant\frac{1}{2}$ but unstable if $\frac{\tau}{h^2}=\frac{1}{2}+\mu$ where $\mu>0$.

In the latter case, the difference scheme (43.57) approximating the first boundary-value problem is also unstable. To prove this assertion, we only need to take $\varphi(mh) = \varepsilon\tilde{X}_{M-1}(m)$, where $\tilde{X}_{M-1}(m)$ is defined by (43.52). The corresponding solution of equations (43.57) is of the form

$$\bar{u}_m^n=\varepsilon(s_{M-1})^n\tilde{X}_{M-1}(m),$$

where

$$s_{M-1}=1-\frac{4\tau}{h^2}\sin^2\frac{(M-1)\pi}{2M}.$$

Since $M=\frac{1}{h}$, we have

$$|s_{M-1}|\longrightarrow\left|1-\frac{4\tau}{h^2}\right|=1+4\mu.$$

as $h\longrightarrow 0$. From this, it is easy to show that the scheme is unstable with respect to the initial conditions.

These examples of the application of the method of separation of variables dealt with two-layer difference schemes. The general idea of the method of separation of variables remains the same for many-layer schemes, but several new phenomena need to be taken into account, which we shall very briefly illustrate with the examples

of two schemes for the wave equation (43.24).

First, let us consider the explicit scheme

$$\frac{\bar{u}_m^{n+1} - 2\bar{u}_m^n + \bar{u}_m^{n-1}}{\tau^2} = \frac{\bar{u}_{m+1}^n - 2\bar{u}_m^n + \bar{u}_{m-1}^n}{h^2}; \qquad (43.58)_1$$

$$\bar{u}_m^0 = \varphi_0(mh); \qquad (43.58)_2$$

$$\frac{\bar{u}_m^1 - \bar{u}_m^0}{\tau} = \varphi_1(mh); \qquad (43.58)_3$$

$$\bar{u}_0^n = \bar{u}_M^n = 0. \qquad (43.58)_4$$

We take as our norm for the right members of the initial conditions $(43.58)_2$ and $(43.58)_3$ the expression

$$\sqrt{(\varphi_0, \varphi_0)_h + (\varphi_1, \varphi_1)_h}.$$

We shall measure the solution in terms of the norm

$$\sup_n \sqrt{(\bar{u}^n, u^n)_h}.$$

Let us first use the method of separation of variables to find a solution of the form $\bar{u}_m^n = T(n)X(m)$ satisfying the boundary conditions $(43.58)_4$. To determine λ and $X(m)$, we again obtain the eigenvalue problem (43.44)-(43.45), the solution of which is given by formulae (43.49)-(43.50). To find the function $T_k(n)$ we have the difference equation

$$T_k(n+1) - \left(2 + \frac{\tau^2}{h^2}\lambda_k\right)T_k(n) + T_k(n-1) = 0. \quad (43.59)$$

We again seek a solution of the boundary-value problem (43.58) in the form

$$\bar{u}_m^n = \sum_{k=1}^{M-1} T_k(n)\tilde{X}_k(m).$$

By using the orthonormality of the system of functions $\{\tilde{X}_k(m)\}$, we obtain from the initial conditions $(43.58)_2$ and (43.58) the initial conditions for $T_k(n)$:

$$T_k(0) = a_k^{(0)}, \quad a_k^{(0)} = (\varphi_0, \tilde{X}_k)_h; \qquad (43.60)$$

$$\frac{T_k(1) - T_k(0)}{\tau} = a_k^{(1)}, \quad a_k^{(1)} = (\varphi_1, \tilde{X}_k)_h. \qquad (43.61)$$

We introduce the fundamental system of solutions

$$\{T_k^{(0)}(n), T_k^{(1)}(n)\}$$

of equation (43.59) defined by the conditions

$$\left.\begin{aligned} T_k^{(0)}(0)=1, \quad \frac{T_k^{(0)}(1)-T_k^{(0)}(0)}{\tau}=0; \\ T_k^{(1)}(0)=0, \quad \frac{T_k^{(1)}(1)-T_k^{(1)}(0)}{\tau}=1. \end{aligned}\right\} \tag{43.62}$$

Then,

$$T_k(n)=a_k^{(0)}T_k^{(0)}(n)+a_k^{(1)}T_k^{(1)}(n),$$

$$\bar{u}_m^n=\sum_{k=1}^{M-1}(a_k^{(0)}T_k^{(0)}(n)+a_k^{(1)}T_k^{(1)}(n))\tilde{X}_k(m).$$

Since the system of functions $\{\tilde{X}_k(m)\}$ is orthonormal, we have

$$(\bar{u}^n,\bar{u}^n)_h=\sum_{k=1}^{M-1}(a_k^{(0)}T_k^{(0)}(n)+a_k^{(1)}T_k^{(1)}(n))^2. \tag{43.63}$$

Let us define

$$P_k(n,\ h)=\max\{|T_k^{(0)}(n)|,\ |T_k^{(1)}(n)|\}$$

and

$$P(n,\ h)=\max_k P_k(n,\ h).$$

From the inequality $(a+b)^2\leqslant 2(a^2+b^2)$ and (43.63), we have

$$(\bar{u}^n,\ \bar{u}^n)_h\leqslant 2\sum_{k=1}^{M-1}(|a_k^{(0)}|^2\,|T_k^{(0)}(n)|^2+|a_k^{(1)}|^2\,|T_k^{(1)}(n)|^2)$$

$$\leqslant 2\,[P(n,\ h)]^2\sum_{k=1}^{M-1}(|a_k^{(0)}|^2+|a_k^{(1)}|^2).$$

On the basis of (43.60) and (43.61), we have

$$\sqrt{(\bar{u}^n,\ \bar{u}^n)_h}\leqslant\sqrt{2}\,P(n,\ h)\sqrt{(\varphi_0,\ \varphi_0)_h+(\varphi_1,\ \varphi_1)_h}. \tag{43.64}$$

It follows from (43.64) that the difference scheme (43.58) is stable with respect to the initial conditions with the choice of norms mentioned above if the quantity $P(n,\ h)$ is bounded for all sufficiently small positive values of h for all n satisfying the inequalities $0\leqslant n\tau\leqslant T$.

On the other hand, it is easy to show that if $P(n,\ h)$ is unbounded, the scheme (43.58) cannot be stable with respect to the initial conditions.

Thus, investigation of the stability of this scheme reduces to investigation of the fundamental system of solutions $\{T_k^{(0)}(n), T_k^{(1)}(n)\}$ of the difference scheme (43.59) that we constructed above.

In the present case, it is comparatively easy for us to make this investigation since $T_k^{(0)}(n)$ and $T_k^{(1)}(n)$ are simply expressed in terms of the roots of the corresponding characteristic equation

$$s_k^2-\left(2+\frac{\tau^2}{h^2}\lambda_k\right)s_k+1=0.$$

We give the results of this investigation: the difference scheme (43.58) is stable with respect to the initial conditions if $\frac{\tau}{h}=c<1$ $(c=\text{const})$ and is unstable if $\frac{\tau}{h}=c\geqslant 1$.

Now, let us consider the implicit difference scheme for the same boundary-value problem

$$\left.\begin{aligned}&\frac{\bar{u}_m^{n+1}-2\bar{u}_m^n+\bar{u}_m^{n-1}}{\tau^2}\\ &=\frac{1}{2}\left(\frac{\bar{u}_{m+1}^{n+1}-2\bar{u}_m^{n+1}+\bar{u}_{m-1}^{n+1}}{h^2}+\frac{u_{m+1}^n-2\bar{u}_m^n+\bar{u}_{m-1}^n}{h^2}\right);\\ &\bar{u}_m^0=\varphi_0(mh);\ \frac{\bar{u}_m^1-\bar{u}_m^0}{\tau}=\varphi_1(mh);\ \bar{u}_0^n=\bar{u}_M^n=0.\end{aligned}\right\}\quad(43.65)$$

This scheme can be investigated in the same way as the scheme (43.58). It turns out that the scheme (43.65) is stable with respect to the initial conditions for arbitrary values of $\frac{\tau}{h}$.

We can use the method of separation of variables to investigate the three-layer difference schemes constructed for the heat-flow equation (43.18) in accordance with the approximating formulae (43.22) and (43.23). In these schemes, the value of the solution on each layer is determined with the aid of the values of the solution on the two preceding layers. The role of the initial conditions is played by the values of the solution on the zeroth and first layers. We write the difference initial conditions defining $\bar{u}_m^0$ and $\bar{u}_m^1$ in accordance with (43.28):

$$\bar{l}_h(\bar{u})\equiv\begin{pmatrix}\bar{u}_m^0\\ \bar{u}_m^1\end{pmatrix}=\begin{pmatrix}\varphi_h(m)\\ \psi_h(m)\end{pmatrix}.\qquad(43.66)$$

The functions φ_h, and ψ_h must be determined in such a way that the difference initial condition (43.66) will approximate the initial condition

$$l(u) \equiv u(0, x) = \varphi_0(x)$$

for the solution of the differential equation.

In accordance with subsection 1, we call the function

$$\beta_h = \begin{pmatrix} \beta^0 \\ \beta^1 \end{pmatrix} = \begin{pmatrix} u(0, mh) - \varphi_h(m) \\ u(\tau, mh) - \psi_h(m) \end{pmatrix}$$

the error approximation for the difference initial condition (43.66). The condition for the approximation to exist consists in the requirement that the approximation error approaches zero as $h \to 0$. In order to give this condition a precise meaning, we need to choose a norm for measuring β_h. For example, we may take the norm

$$\sqrt{(\beta^0, \beta^0)_h + (\beta^1, \beta^1)_h}.$$

This norm can also be used for measuring the difference initial conditions. As in the preceding examples of this paragraph, the change in the solution will be measured in terms of the norm

$$\sup_n \sqrt{(\bar{u}^n, \bar{u}^n)_h}.$$

With this choice of norms, it is easy to obtain the following results. The explicit difference scheme

$$\frac{\bar{u}_m^{n+1} - \bar{u}_m^{n-1}}{2\tau} - \frac{\bar{u}_{m+1}^n - 2\bar{u}_m^n + \bar{u}_{m-1}^n}{h^2} = 0,$$

$$\bar{u}_m^0 = \varphi_h(m), \quad \bar{u}_m^1 = \psi_h(m), \quad \bar{u}_0^n = \bar{u}_M^n = 0$$

is unstable with respect to the initial conditions for all values of $\frac{\tau}{h^2} = \text{const}$. The implicit difference scheme

$$\frac{3}{2}\frac{\bar{u}_m^{n+1} - \bar{u}_m^n}{\tau} - \frac{1}{2}\frac{\bar{u}_m^n - \bar{u}_m^{n-1}}{\tau} - \frac{\bar{u}_{m+1}^{n+1} - 2\bar{u}_m^{n+1} + \bar{u}_{m-1}^{n+1}}{h^2} = 0,$$

$$\bar{u}_m^0 = \varphi_h(m), \quad \bar{u}_m^1 = \psi_h(m), \quad \bar{u}_0^n = \bar{u}_M^n = 0$$

is stable to the initial conditions for arbitrary values of $\frac{\tau}{h^2}$.

6. We can apply the Fourier-integral method to investigate the stability with respect to the initial conditions in the case of the Cauchy problem. This method is convenient for equations whose coefficients do not depend on the variable x (in particular, for equations with constant coefficients). Just as with the method of separation of variables, the Fourier-integral method makes it possible to reduce the investigation of stability to a study of the solutions of some 'ordinary' difference equation (that is, a difference equation containing only one variable index).

The basic idea of the method consists in investigating the solutions of the difference scheme that correspond to the initial functions of the form e^{ikx}, where k is a real parameter. With the aid of such solutions, we immediately obtain necessary conditions for stability. By using the Fourier integral*, which enables us to express arbitrary initial conditions in terms of functions of the form e^{ikx}, we can also obtain sufficient conditions for stability. Here, we confine ourselves to a derivation of the basic necessary condition for stability in the case of two-layer schemes.

We shall call the function $e^{ikx} = e^{ikmh}$. Let us try to find a solution of the difference scheme corresponding to the initial function $\bar{u}_m^0 = e^{ikmh}$ in the form

$$\bar{u}_m^n = T(n, k, h)\, e^{ikmh}. \tag{43.67}$$

To determine the function $T(n, k, h)$ we obtain a difference equation involving the variable n. If, for a given $k = k_0$ the function $T(n, k, h)$ will be bounded for sufficiently small h and all n such that $n\tau \leqslant T_0 = \text{const}$, this means that the scheme is stable on the harmonic e^{ik_0x}. If $T(n, k, h)$ is uniformly bounded for all k under the conditions imposed on h and n, this means that the difference scheme is uniformly stable on all harmonics.

With the norms $\|\bar{u}^0\| = \sup\limits_m |\bar{u}_m^0|$ and $\|\bar{u}_h\| = \sup\limits_{m,\,n} |\bar{u}_m^n|$ for the scheme to be stable with respect to the initial conditions, it is obviously necessary that it be uniformly stable on all harmonics. It is fairly easy to test whether this condition is satisfied, and, for that reason, it is widely used in practice for investigating difference schemes.

* The reader may acquaint himself with the Fourier integral by consulting SHILOV, G.E., *Mathematical analysis, Special course*, Fizmatgiz, Chapter VII (1960).

Let us consider some simple examples. Let us find $T(n, k, h)$ for the difference scheme (43.32). When we substitute (43.67) into (43.32), we obtain

$$T(n, k, h) = [s(k, h)]^n, \text{ where } s(k, h) = 1 + \frac{\tau}{h} - \frac{\tau}{h} e^{ikh}.$$

Let us show that the condition for uniform boundedness is not satisfied for $T(n, k, h)$ for any constant value of $\frac{\tau}{h}$. We set $kh = \alpha$. Then,

$$|s|^2 = \left[1 + \frac{\tau}{h}(1 - \cos\alpha)\right]^2 + \frac{\tau^2}{h^2}\sin^2\alpha.$$

For $\alpha = \frac{\pi}{2}$, we have

$$|s|^2 = \left(1 + \frac{\tau}{h}\right)^2 + \frac{\tau^2}{h^2}.$$

Therefore, the values of $T(n, k, h)$ are not bounded and the scheme (43.32) is not stable with respect to the initial conditions for any value of $\frac{\tau}{h}$.

Now, let us consider the scheme (43.33). In the case of this scheme, $T(n, k, h) = [s(k, h)]^n$, where

$$s(k, h) = 1 - \frac{\tau}{h} + \frac{\tau}{h} e^{ikh}.$$

From this, we get

$$|s|^2 = \left[1 - \frac{\tau}{h}(1 - \cos\alpha)\right]^2 + \frac{\tau^2}{h^2}\sin^2\alpha. \qquad (43.68)$$

If $\frac{\tau}{h} = c \leqslant 1$, it follows from (43.68) that $|s(k, h)| \leqslant 1$ and, consequently, the values of $T(n, k, h)$ are uniformly bounded. However, if $\frac{\tau}{h} = c > 1$, then, for $\alpha = \pi$ we have

$$|s|^2 = (1 - 2c)^2 > 1 + \mu,$$

where $\mu > 0$ does not depend on h. For the corresponding value of k, the function $T(n, k, h)$ increases in absolute value without bound as $n \to \infty$. Therefore, the scheme (43.33) is unstable with respect to the initial conditions when $c > 1$.

Let us now consider an implicit difference scheme for the same problem (43.31) constructed according to (43.16) (a second-order approximation error with respect to τ and h).

For this scheme,

$$s(k,\ h)=\frac{1-i\frac{\tau}{h}\operatorname{tg}\frac{\alpha}{2}}{1+i\frac{\tau}{h}\operatorname{tg}\frac{\alpha}{2}}.$$

as is easily verified. Therefore, $|s|=1$ and the values of $T(n,\ k,\ h)$ are uniformly bounded for an arbitrary value of $c=\frac{\tau}{h}$.

By a similar procedure, we can also investigate three-layer schemes.

7. The concept of a correct difference scheme is connected not only with the question of the convergence of the solutions of the difference schemes but also with the question (very important in practice) of the effect of rounding-off errors on the approximate solution obtained by means of the difference scheme. In practice, all calculations are made with rounding-off, which has some effect or other on the solution of the difference scheme. Obviously, only those schemes are of practical interest for which the small errors incurred in the process of numerical solutions of the difference equations do not lead to significant deviations from the exact solution of these equations.

We shall give some simple examples that illustrate the theoretical distinction between correct and incorrect difference schemes from the point of view of increase in the error incurred by rounding-off. Let us look at the difference schemes corresponding to the first boundary-value problem for the heat-flow equation (43.18).

Let us give an estimate of the rounding-off error for the explicit difference scheme constructed in accordance with (43.39) for $\frac{\tau}{h^2}\leqslant\frac{1}{2}$. Suppose that the error made on each individual layer as a result of rounding-off in the calculation of $\overline{u}_m^n$ does not exceed ε in absolute value. Let us first assume that this error is made only on the n_0th layer and that no rounding-off errors are made on any of the other layers. For the scheme in question with $\frac{\tau}{h^2}\leqslant\frac{1}{2}$ the error ε admitted on the n_0th layer causes a change from the exact solution of the difference scheme (on the following layers)

that does not exceed ε. This follows immediately from equation (43.39) since, in view of the linearity of the difference equations, the difference between the exact and approximate solutions is also a solution of equation (43.39) for $n > n_0$.

To get an estimate of the rounding-off error in the general case in which errors in rounding-off are made in all layers, it will be sufficient, by virtue of the linearity of the difference equations, to add the errors in rounding-off that are made on the individual layers. For $t \leqslant T$, the number of layers is bounded by the quantity $\frac{T}{\tau}$. Therefore, the total error in rounding-off does not exceed $\frac{\varepsilon T}{\tau}$ in absolute value.

The rounding-off error will not increase as $h \to 0$ if $\varepsilon = O(\tau) = O(h^2)$. Let us denote by ρ the maximum rounding-off error that is incurred in carrying out the elementary arithmetic operations performed in evaluation $\overline{u}_m^{n+1}$ from formula (43.39). It follows from this formula that $\varepsilon = O(\rho)$ Therefore, when $\rho = O(h^2)$ the rounding-off error will be bounded as $h \to 0$. (We note that, for the implicit scheme constructed from formula (43.20), we obtain the condition $\varepsilon = O(h^2)$, in exactly the same way, though the error ε in terms of ρ has a more complicated form. This estimate depends on the chosen method of solving the system of algebraic equations connecting the values of $\overline{u}_m^{n+1}$ with the $(n+1)$st layer.)

Suppose now that we have

$$\frac{\tau}{h^2} = \frac{1}{2} + \mu$$

with $\mu > 0$ in the scheme constructed from (43.39). (This scheme is unstable with respect to the initial conditions.) As was shown in subsection 4, an error of the form $(-1)^m \varepsilon$ made on the initial layer will cause an error on the $n = \frac{t}{\tau}$ th layer (where $t \leqslant T$ that is equal in absolute value to the quantity $\varepsilon e^{\frac{Kt}{h^2}}$, where K is a constant. For the rounding-off error not to increase as $h \to 0$, ε must decrease extremely fast as $h \to 0$, specifically, as fast as $e^{-\frac{Kt}{h^2}}$

We now give a numerical example illustrating a possible rapid increase in the errors in unstable schemes. Suppose that $\frac{\tau}{h^2}$ is equal to unity in the scheme constructed from

(43.39). Then, an error of the form $(-1)^m\varepsilon$ increases three times as we go from n to $n+1$. After twenty steps, it increases $3^{20} \approx 3.5 \cdot 10^9$ times. If the calculations are performed with a relative accuracy of 10^{-9}, the results obtained on the 20th layer will not, in general, contain a single accurate digit.

The question of the accumulation of rounded-off errors that are made in the calculation of the solution of difference equations has as yet received little study.

8. The examples that we have been considering of the simplest boundary-value problems for differential equations and the difference schemes corresponding to them give only a general idea of the basic concepts associated with the method of nets and the most important methods of investigating difference schemes.

A large number of studies have now been made on different general questions dealing with the theory of difference schemes. Certain particular classes of schemes have been studied and individual schemes that are important for practical applications have been investigated in detail*.

The investigation of the stability of difference schemes occupies the leading place in these studies. As a rule, it is difficult to obtain effective (that is, more or less easily verified) sufficient conditions for convergence. It is usually comparatively easy to find necessary conditions. In practice, great significance is attached to simple and at the same time strong (that is, almost sufficient) necessary conditions for stability. There are methods that make it possible to obtain such conditions for certain rather general classes of schemes (for example, the Fourier-integral method for difference schemes with constant coefficients, which was briefly expounded above).

However, it should be noted that the difference schemes that are applied in practice do not in the majority of cases admit a complete rigorous investigation with the aid of general methods that are known at the present time. In particular, this applies to schemes corresponding to linear differential equations with variable coefficients of a general form. The studies that have been made on the stability of

* See RICHTMYER, R.D., *Difference methods for initial value problems*, New York, Interscience (1957).

such schemes use specific properties of each particular scheme and they cannot be used as models for investigating the various types of schemes that are used in actual computation.

Also, we encounter cases in which the application of general methods of investigation is theoretically possible but which involve such great theoretical difficulties that these methods are not suitable in practice. (For example, considerable difficulties in calculating eigenvalues may be encountered in the investigation of stability by the method of separation of variables.)

Therefore, so-called practical techniques of investigating the stability of difference schemes have been found and have been widely used. Theoretically, these techniques have not been rigorously justified or have been justified only for particular cases but they have been quite well verified in practice.

One of these techniques is the so-called method of freezing coefficients. According to this method, linear difference equations with variable coefficients are replaced with the same equations with constant coefficients equal to values of the corresponding variable coefficients at some point (t_0, x_0) belonging to the region in question. If, for an arbitrary choice of points (t_0, x_0), the scheme consisting of the equations with constant coefficients turns out to be stable, the original scheme (with variable coefficients) is assumed to be stable.

There are also a number of other practical techniques of investigating the stability of difference schemes.

Clarification of the limits of applicability and rigorous justification of these practical techniques of investigation of stability of difference schemes are of considerable interest.

Along with the development of general methods of investigating difference schemes, the development of rational methods of construction of difference schemes for certain classes of problems that often arise in applications is of great importance. Of especial interest are difference schemes that are suitable for finding approximations of discontinuous and nonsmooth solutions of linear and (especially) nonlinear differential equations.

Questions that are interesting from a theoretical standpoint and important in practice arise in the numerical solution of partial differential equations in the case in which the

number of independent variables exceeds two. Ordinary difference schemes analogous to those that we considered above lead in this case to a net consisting of a very great number of nodes. Here, the amount of calculation increases sharply and so does the amount of information that we need to store in the memory devices of a computing machine. Therefore, exact estimates of the information necessary for obtaining a solution of prestated accuracy are of great importance for the development of methods of approximate solution of many-dimensional problems, and new methods of constructing perfected difference schemes making it possible to decrease the number of computations are of great importance.

Index